DOCUMENT RESUME

ED 215 885 SE 037 079

AUTHOR	Suppes, Patrick; And Others
TITLE	Research on Process Models of Basic Arithmetic Skills, Technical Report No. 303. Psychology and Education Series - Final Report.
INSTITUTION	Stanford Univ., Calif. Inst. for Mathematical Studies in Social Science.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.; National Science Foundation, Washington, D.C.
PUB DATE	1 Sep 81
GRANT	SED-78-22286
NOTE	99p.; NIE-NSF Joint Program for Research on Cognitive Process and the Structure of Knowledge in Science and Mathematics.
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Addition; *Algorithms; Cognitive Processes; Computation; Computers; Educational Research; Elementary Education; *Elementary School Mathematics; *Eye Movements; Mathematics Education; Mathematics Instruction; Models; *Subtraction
IDENTIFIERS	*Mathematics Education Research

ABSTRACT

This report presents a theory of eye movement that accounts for main features of the stochastic behavior of eye-fixation durations and direction of movement of saccades in the process of solving arithmetic exercises of addition and subtraction. The best-fitting distribution of fixation durations with a relatively simple theoretical justification consists of a mixture of an exponential distribution and the convolution of two exponential distributions. The eye movements themselves were found to approximate a random walk that fit rather closely in both adult and juvenile subjects; the motion postulated by the normative algorithm ordinarily taught in schools. Certain structural features of addition and subtraction exercises, such as the number of columns, and the presence or absence of regrouping, are well known to affect their difficulty. In this study, regressions on such structural variables were found to account for only a relatively small part of the variation in eye-fixation durations. (Author/MP)

ED215885

Research on Process Models of Basic Arithmetic Skills

by

Patrick Suppes, Michael Cohen, Robert Laddaga,

James Anliker, and Robert Floyd

Technical Report No. 303

September 1, 1981

Psychology and Education Series

Final Report

National Institute of Education-National Science Foundation
Joint Program for Research on Cognitive Process
and the Structure of Knowledge in Science and Mathematics

Grant Number: SED78-22286

Institute for Mathematical Studies in the Social Sciences
Stanford University
Stanford, California

Summary

This report presents a theory of eye movement that accounts for main features of the stochastic behavior of eye-fixation durations and direction of movement of saccades in the process of solving arithmetic exercises of addition and subtraction. The best-fitting distribution of fixation durations with a relatively simple theoretical justification consists of a mixture of an exponential distribution and the convolution of two exponential distributions. The eye movements themselves were found to approximate a random walk that fit rather closely in both adult and juvenile subjects the motion postulated by the normative algorithm ordinarily taught in schools. Certain structural features of addition and subtraction exercises, such as the number of columns, the presence or absence of a carry or a borrow, are well known to affect their difficulty. In our study, regressions on such structural variables were found to account for only a relatively small part of the variation in eye-fixation durations.

Table of Contents

Acknowledgments

The authors would like to acknowledge the assistance of George Black and Tom Dienstbier for the design and implementation of the DART and the keyboard and for computer maintenance. We thank William R. Sanders and Ron Roberts for programming and system-design work. Barbara Anderson has been helpful in many ways, particularly with taping and organizing data files.

1 Introduction and Theoretical Background

Since the second decade of this century there has been a large number of empirical studies, many excellent, of student performance in arithmetic. A fairly systematic review of this early literature is to be found in Suppes, Jerman, and Brian (1968). More recently, linear regression models that use the structural features of the arithmetic exercises as independent variables have been shown to fit the data of responses or of response latencies rather well (see Suppes and Morningstar, 1972, for extensive experimental data and references to previous literature. The move from the structural regression models to process models that are finite-state automata is relatively straightforward. From a certain formal standpoint, automaton models are algorithmically sufficient to the task. They provide a process analysis of the steps needed to solve the standard exercises (see Suppes, 1969, for a theoretical development of such automaton models; see Suppes and Morningstar, 1972, Chapter 5, for the fit to data). The automaton models are, however, at too abstract a level to be fully satisfactory from a psychological standpoint, even though they are mathematically satisfactory as algorithms.

The next step beyond automaton models in the line of research reported here is to introduce schematic concepts of perceptual processing. Because the theory behind register machines with perceptual instructions is a discrete version of the stochastic and continuous model involving eye movements that we consider later, we turn to some of the details of these register-machine models, which were first proposed in Suppes (1973).

First of all, we emphasize again the importance of a perceptual component. Without such a component we have no way of representing the mental operations a person actually uses to process the written symbols presented to him. It is also apparent, on the other hand, that the full theory of how this processing takes place is a topic of unbounded complexity. We will necessarily provide a treatment at a certain level of abstraction but one that is closer to the concrete and complex aspects of the actual perceptual situation than is that of the automaton models.

Register machines as such were first introduced by Shepherdson and Sturgis (1963) in order to give a natural representation of computable functions. The representation is more intuitive than that of Turing machines because the central idea is close to that of a standard computer accepting instructions. For the representation of computable functions, a rather small set of arithmetic instructions is sufficient. In particular, an unlimited register machine has a denumerable sequence of registers, but any given program only uses a finite number of these registers. The Shepherdson and Sturgis machine accepts six basic instructions: add one to a register, subtract one, clear a register, copy from one register to another, and two jump instructions, one conditional and one not. (It is apparent that this set of instructions is not minimal; the exact choice is more or less a matter of convenience.)

To model the processing that a person does, we want a different register machine, of course, and a quite different set of instructions. In particular, we want instructions that reflect the perceptual situation. It is also reasonable to assume that only a fixed finite number of registers are used in the relatively simple tasks we consider.

We can drastically simplify the perceptual situation by conceiving of each exercise, for example, an exercise in column addition, as being presented on a grid with at most one symbol in each cell of the grid. The person doing exercises is represented by a model that has instructions for attending to a specified cell; for example, in the standard algorithms of addition, subtraction, and multiplication, most of us were taught to begin at the upper right-hand corner and to move downward through each column and from right to left across columns. We shall discuss the detailed set of instructions in a moment. The basic idea of such register-machine models is that the different algorithms of arithmetic are represented by subroutines. One subroutine may be called in another as complex routines or programs are built up. For instance, the routine for column multiplication uses as a subroutine the program for column addition.

To have a psychologically realistic model at this level, the problem is to find a representation that is not only adequate as an algorithm but that also can be fitted to detailed eye-movement data in the same way that the linear regression models or the automaton models mentioned earlier have been applied to response and latency data.

1.1 Register-machine Model for One-column Addition

To avoid some complexities that we need to consider later in terms of the actual experiments we performed, we consider initially the simple case of one-column addition. Contrary to the usual Cartesian convention, we number the coordinates of the grid on which we think of symbols being presented from the upper right-hand corner. Thus, in the addition exercise

$$\begin{array}{c}6\\3\\\underline{9}\end{array},$$

the coordinates of the digit 6 are (1,1), the coordinates of 3 are (2,1), and the coordinates of 9 are (3,1), with the first coordinate being the row number and the second being the column number (the column number is needed for the general case).

In formulating instructions we need for column addition, the following notation is used: (a,b) is the pair of coordinates of a grid position with a and b being positive integers; (±a, ±b) shows the amount of shift in each coordinate from one grid position to another; [R], $[R_1]$ and $[R_2]$ are variables for registers; [SS] is the fixed register that is stimulus supported; [NSS] is a fixed register that is nonstimulus supported. These are the instructions:

Attend (a,b):	Direct attention to grid position (a,b).
(±a, ±b):	Shift attention on the grid by (±a, ±b).
Readin [SS]:	Read into the stimulus-supported register the visual symbol in the grid position addressed by Attend.
Lookup [R1] + [R2]:	Look up table of basic addition facts for adding contents of registers [R1] and [R2] and store the result in [R1].
Copy [R1] in [R2]:	Copy the content of register [R1] into register [R2].
Jump (val) R,L:	Jump to line labeled L if content of register [R] is val.
Outright [R]:	Write (output) the rightmost symbol of register [R] at grid position addressed by Attend.
Deleteright [R]:	Delete the rightmost symbol of register [R].
End:	Terminate processing of current exercise.

Of these instructions, only Lookup does not have an elementary character. In a more complete analysis, it would have the status of a subroutine built up from more primitive operations such as those of counting. It is, of course, more than a problem of constructing the table of basic addition facts from counting subroutines; it is also a matter of being able to add a single digit to any number stored in the nonstimulus-supported register [NSS], as, for example, in adding many rows of digits in a given column. We omit the details of building up this subroutine. It should also be obvious that the remaining instructions do not constitute a minimal set.

For the simple case of one-column addition, we need only two registers; one, [SS], is stimulus-supported; the other, [NSS], is not. A program representing a procedure close to what is taught in schools for doing one-column addition is shown in the left-hand column of Table 1. A restriction is that the sum be equal to or less than 99. By adding probabilistic parameters to various segments of the program, response performance models are easily generated. The more complex routines required for general addition and subtraction exercises are given in Appendix C.

Because it is doubtful that a young student could be taught routines stated in terms of the assembly language we use in our set of instructions, in the right-hand column of Table 1 we have written down an English version of the program and have called the various commands English-addressable subroutines (for an amplification of this point, see Suppes, 1980). It is not our intention here to enter into those aspects of learning for which the English-addressable subroutines are nearly essential. We do think that a very sensible and reasonable cognitive theory of learning for such procedures can be formulated by using such subroutines, but in the present report we shall restrict ourselves to performance models and shall not consider in any systematic way how the performance models have been learned.

Table 1

Example of One-column Addition for Sums < 99

	Pseudo assembly-language program	English-addressable subroutines
	Attend (1,1)	Look at this number.
	Readin	
	Copy SS in NSS	Remember the number.
	Attend (+1,+0)	Now look at this next
	Readin	number.
Opr	Lookup NSS + SS	Add the two numbers and remember the sum.
	Attend (+1,+0)	Move down the column.
	Readin	If there is another number,
	Jump (0-9) SS, Opr	add as before and continue.
	Attend (+1,+0)	If not, move down to the blank space.
	Outright NSS	Write down the number of ones in the answer.
	Deleteright NSS	
	Attend (+0,+1)	Now look at the space to the left.
	Outright NSS	Write down the number of tens in the answer (unless it is zero).
	Deleteright NSS	
	End	

1.2 Normative Versus Actual Performance

At the level of the register-machine models discussed above and as exemplified in Table 1 and Appendix C, we have a clear concept of the normative behavior that every good teacher aims at in the elementary school. Students first learn subroutines of the kind we have been discussing and then they are expected to work standard exercises with few, if any, errors. We think of such accuracy as characteristic of the normative model. Of course, we do not have such clear ideas about the response latencies to be expected but there are qualitative norms, implicit in classroom practice, about working at an acceptable speed. The empirical and experimental studies of arithmetic mentioned earlier focus on actual performance and how it deviates from normative standards.

The situation is very different once we introduce eye movements. There is no obvious or natural concept of normative eye movements and as far as we know there has been no discussion in the literature of what one would take to be appropriate eye movements as a real-time process on the part of children or adults. We return to say something about such normative behavior later, one of the reasons being that we have considerable data from skilled adults, but our main concern is the following.

1.3 Stochastic Theory of Eye Movements

To move toward a much more detailed theory of how programs of the kind exemplified in Table 1 are actually executed we move now to theory and data on eye movements. We expect a strong correlation between the point of regard and the current step in a procedure that is being executed. We recognize there are many exceptions to this claim, for example, when someone looks off into the distance while recalling some past event or planning some future activity. But there are just as many cases in which this is obviously not so and the point of regard in the field of visual perception is an obvious clue to what is being processed internally. In this particular we are advancing the view that the point of regard within the displayed arithmetic exercise, taken in conjunction with a knowledge of the algorithm which the subject has been instructed to execute, provides an important clue to what covert mental operations the subject is performing and where he is at that moment in the arithmetic algorithm.

We turn now to the first simple version of our stochastic theory of eye movements. It is to be understood that the theory is formulated for the context of performing algorithmic tasks represented by subroutines such as shown in Table 1. Perhaps the conceptual point of widest psychological interest is the contrast between the stochastic theory of human behavior at the level of eye movements and the essentially deterministic theory of algorithmic processing in standard computer systems. It is possible to build stochastic features into standard computers but it is seldom done except for purposes of simulation, whereas such features are a fundamental aspect, in our judgment, of human execution of procedures. This highly stochastic aspect of human performance has not been adequately incorporated into current cognitive theories of behavior, but is widely recognized in the literature on eye movements (Ditchburn, 1981).

Our simple stochastic theory of eye movements falls naturally into two parts. The first part is the simple qualitative axiom governing the duration of fixation. We comment later on its psychological significance. The second group of axioms has to do with the direction of eye movements. In the present context we think of the direction of movement as constituting a random walk, with the frame of reference of the steps being the cognitive grid defined earlier as implicit in the perception of the stimulus displays of arithmetic exercises. The most important random-walk movement is going to the next displayed symbol. We call this the forward jump, because it represents progress in the algorithmic sequence. The second significant movement is that of staying within a given square on the grid. We call this the stayput movement, which is similar to the concept of gaze in Just and Carpenter (1980). Notice that the point of regard is not literally 'staying put' at the level of resolution of the eye-tracking apparatus but rather is doing so in terms of the large-scale grid on which stimuli are displayed. The point of regard stays within the same square of the grid, i.e., stays focused on approximately the same displayed symbol. The third type of movement is that of backtracking, which is the opposite of forward movement. It is to be noted that backtracking as such will not occur in the normative discrete routines as exemplified in Table 1 or Appendix C.

1.4 Model I

The first axiom, F1, is unconditional in character about the length of fixation, showing that the sense of a conditional action is not needed.

Simple Axiom of Fixation

F1. The length of fixation is independent of past (cognitive) processing and the present stimulus context.

It follows at once from this axiom by standard probabilistic arguments that the distribution of fixation lengths is exponential. What is important from a psychological standpoint is that the axiom implies that the length of fixation is a process without memory. What the organism is currently doing or what the organism has recently been doing has no effect at all on the length of fixation. This is a very strong independence assumption and we shall want to examine in detail the extent to which it is satisfied by our data.

Axioms on Direction of Movement

D1. If processing is complete at a given point of regard, then jump to the next stimulus symbol.

D2. If processing at a given point of regard is not complete and nonstimulus-supported memory has not decayed, stayput on present stimulus symbol.

D3. If processing at the present point of regard is not complete and nonstimulus-supported memory has decayed, backtrack to the beginning of the exercise.

It is apparent how these three axioms are tied to the use of the nonstimulus-support registers described earlier and exemplified in the routine shown in Table 1. For some of the more complicated algorithms we consider (see Appendix C), it is not sufficient to have simply the two registers referred to in Table 1 but it is necessary to have several, and what we say in these axioms about nonstimulus-supported memory is meant to apply to any of the registers that is not stimulus-supported.

We emphasize that Model I is meant to be a natural expansion of the discrete register-machine model, of which a simple example has been given in Table 1. The axiom on fixation times and the axioms on direction of movement define in a natural way a simple continuous-time process. The random walk that is part of this process is a little unusual because of the strong backtracking assumption but the axioms taken together do define what seems to be the obvious continuous-time Markov process to add eye-movement phenomena to the original register-machine model.

1.5 Model II

As we shall see in the data presented below, there are at least four major respects in which Model I is too simple. First, the distribution of fixation times is not exponential, which means that the process is not completely without memory. Furthermore, if we reflect on this problem it is evident that even from a physical standpoint we would expect some kind of inertial effect that would prohibit the process from being strictly exponential. Second, the backtracking axiom is certainly too simple in formulation. As we shall see, the backtracking observed in our subjects is only one or two grid spaces back and seldom a full backtrack to the beginning of the exercise. The implication of this is strong for the formulation of the register-machine model. It means that the number of registers used, for example in addition or subtraction, must be increased, in order to store partial results, or the character of the registers must be changed. These two approaches, as can easily be seen, are formally equivalent. Third, there is good evidence in the data that a certain percentage of the time a square on the grid is skipped. This skipping phenomenon may be accounted for by the subject's being able to identify at one point of regard more than one stimulus symbol. We shall not enter here into the question of whether or not it is accomplished by peripheral vision; but the skipping phenomenon is certainly present and needs to be incorporated into an adequate theoretical model. Fourth, there are instances in which the algorithmic routine has not been followed, and backtracking has not occurred. In the extended random walk of Model II, we call this step <u>other</u>. It might be thought that any movements like <u>other</u> outside the theory are to be regarded as evidence against the theory but we take it that with something as stochastic and irregular as eye movements we can scarcely expect to have a completely detailed account of the movement, and we need this category to describe fully the eye motion. It must be admitted that from a theoretical standpoint we still have a rather poor understanding of how to think about these 'other' movements.

We take these four points into consideration in revising the axioms of Model I to characterize Model II. First, in the case of fixation duration we assume that a fixation is exponentially distributed when a single instruction is being executed internally and is the convolution of n exponentials (with the same parameter) when n internal instructions are being executed during that fixation. (These internal instructions are thought of as controlling eye movement in fine detail, and are not the kind of instructions used in Table 1.) In the data analysis presented later we consider only the case of $n = 2$, but it is evident that we could obtain a better fit by increasing n. The concept of the number of instructions being executed internally is a theoretical one, since we can make no direct observation of n. Those who do not like this concept of an instruction being executed can easily supply an alternative, more phenomenological, formulation.

<u>Mixture Axioms on Execution Time</u>

F1. <u>The execution of each eye-control instruction is independent of past processing and the present stimulus context</u>.

F2. <u>Each fixation lasts for the execution of</u> n <u>internal instructions</u>, <u>for</u> $n = 1, 2, \ldots$.

These two axioms are too weak by themselves to imply an interesting parametric form for the distribution of fixation durations. For the purposes of present application to our data, we specialize to $n = 1$ or $n = 2$. Moreover, the exponential parameter for the convolution of two executions we permit to be different than that for one. Given α as the mixture parameter for the exponential distribution with mean λ_1 and λ_2 as the mean of the convolution of two exponentials, the distribution of fixation durations for Model II has the following form:

$$f(t) = \frac{\alpha}{\lambda_1} e^{-t/\lambda_1} + \frac{(1-\alpha)t}{\lambda_2^2} e^{-t/\lambda_2}$$

We now turn to the revision of the axioms on the direction of eye movement. The most important revision is in the axiom regarding backtracking. The simple normative model is almost never exhibited in the data, i.e., there was a vanishing small probability of backtracking to the origin point of the grid, except in the case when the subject was in the first column and this was covered by the classification given below. Most of the backtracking was only to the preceding row in the same column. We therefore postulate that backtracking is only to the preceding row or at most to the beginning of the column, i.e., to the top of the column. Of course, in subtraction exercises these two forms of backtracking are identical and they are also identical in many forms of addition. We take account of this fact in the data analysis given later. We also add an axiom to take account of skipping, which can scarcely be regarded as a deviation from execution of the algorithm but represents a feature not represented in Model I. It should be mentioned that except with negligible probability the skipping always involved just one square of the grid, i.e., ordinarily one symbol.

Revised Axioms on Direction of Movement

Axioms D1 and D2 remain unchanged.

Axiom D3′. If processing at the present point of regard is not complete and nonstimulus-supported memory has decayed, backtrack to the immediately preceding row in the same column or to the beginning of the column.

Axiom D4′. If the present point of regard also provides a perceptual image of the next symbol and processing, it is complete; then skip over the next stimulus symbol to the following one.

It is apparent that the modified backtracking axiom requires the introduction of complications in the registers in which perceptual data are stored. It is also clear, as already remarked, that this can be accomplished in a number of ways. Perhaps the simplest way in the case of addition is to keep not only the current partial sum but the preceding partial sum in the nonstimulus-supported register. If we want, this can also be done by simply adding another register. One of the registers keeps the current partial sum and the other the partial sum that preceded it. This takes care of backtracking one row. When backtracking is to the beginning of the column we also need to have somewhere storage of the carry, in the case of addition, or borrow, in the case of subtraction, and for this purpose still another register is easily added. In Appendix C we give the full procedures in terms of the normative models for multi-column addition and multi-column subtraction, but we decided not to add a still more complicated setup for these additional registers. In the case of backtracking we include data analysis, and we believe it is obvious enough how these additional registers can be added to accommodate the features required by the axiom on backtracking. Concerning the detailed algorithms, the same remarks apply to the axiom on skipping.

As we shall see, the fit of Model II to the data is certainly not perfect, but it represents the main features of the data fairly well. We shall not consider in this article any major extensions but we do discuss in the last section matters that seem worth further investigation.

1.6 Related Theoretical Work

The detailed procedural theory we have proposed for doing arithmetic exercises does not seem to have a close predecessor in the literature, but there are related proposals for different tasks. We mention especially Groner's work (1978); he proposes various Markov models for eye movements in such standard tasks as the three-series problem, but the details are quite different from what we have proposed. Closer in spirit is the recent work of Just and Carpenter (1980) on a theory of eye fixations in reading comprehension.

Their theory rests on two general assumptions. The first is that of immediacy, i.e., that a reader tries to interpret each content word of a text as it is encountered. It is apparent that our procedures for addition and subtraction satisfy this assumption: For example, in adding many rows a partial sum is stored, not the full sequence of numbers attended to. Their second assumption is that the eye remains fixated on a word as long as the word is being processed. Just and Carpenter call this the eye-mind assumption. It corresponds rather closely to Axiom D2 of Models I and II.

Just and Carpenter go on to consider in a qualitative way the evidently large number of processing stages required for reading--processes such as encoding, lexical access, semantic interpretation, and appraisal of context. It is easy to see from this list or its variants that the process of reading text is a much more complicated task than the one we have studied of performing arithmetic algorithms. There is a close relation between their two general assumptions and some of our axioms, but we part on details because of the different tasks studied. Due to the greater simplicity of arithmetic tasks we have been able to present a more detailed and complete theory than they have for reading, and consequently to study rather thoroughly the quantitative aspects of the theory we have proposed.

2 Method

In this section we will first give a brief account of the methods employed in designing and performing the experiments reported in this article. First we will cover the methods for the experiment with adult subjects, and then indicate in what ways the method for the experiment with children differed. Following that, we have provided a somewhat detailed account of a preliminary study of some calibration issues.

2.1 Method for adult subjects

The study of adult subjects served two purposes. One was to obtain a benchmark concerning the eye movements of expert subjects. The second was to test the effectiveness of the algorithm used in the initial register-machine model as a normative model for human eye-movement behavior. In this case the subjects, who were two adult male college students, were informed about the algorithms and asked to try to follow them, including the appropriate eye movements, in their computations.

Since we were testing normative aspects of the register-machine model using expert subjects, a simple randomized procedure was used to generate the arithmetic exercises used in the experiment. Digits for each position in each exercise were preselected randomly, subject only to the following constraints: (a) no left digits may be zeros, (b) each of the addition exercises must have an equal number of rows and columns (e.g., 3 X 3, 4 X 4, etc.); (c) in the addition exercises, no two horizontally adjacent digits can be the same, and no digit can be repeated in a column; (d) in the subtraction exercises, zero differences were not permitted in the leftmost column.

There were 720 addition exercises and 600 subtraction exercises used in the adult studies. These exercises were divided into 12 sessions of addition exercises and 10 sessions of subtraction exercises, with 60 exercises in a session. Each addition session included 20 exercises with 3 columns and 3 rows of digits, 20 exercises with 4 columns and 4 rows of digits, 10 exercises with 5 columns and 5 rows of digits, and 10 exercises with 6 columns and 6 rows of digits.

2.2 Equipment and Setup

The experimental apparatus included a computerized eye-tracking system, a display terminal, and a simplified keyboard for the subject's manual responses. The computer-based eye-tracking system, known as PERSEUS (Anliker, et al., 1977) incorporates as a peripheral hardware device a 2-dimensional double-Purkinje-image eye-tracker (DPIET) developed by Cornsweet and Crane (1973) and updated by Crane and Steele (unpublished, 1977). PERSEUS uses advanced software--implemented on a medium-sized computer (PDP-15/76)--(a) to calibrate each subject's eye-pointing responses, (b) to correct for linear and nonlinear systematic eye-tracking errors, (c) to detect, measure and record higher order phenomena (e.g., fixations, saccades, and scanpaths) and (d) to control the real-time aspects of stimulus presentation and data collection. PERSEUS delivers highly accurate (less than 5 minutes of arc error over a field of 20 X 20 degrees) measurements of a subject's point-of-regard.

The PERSEUS system was interfaced with a dual DEC KI10 system running a modified version of the TENEX operating system. The KI10 system was used to select and present pre-compiled exercise items. Both PERSEUS and the KI10 system were interfaced with an IMLAC PDS-1 minicomputer display system, and a subject keyboard. During experimental sessions all communications from the keyboard to the KI10 and from the KI10 to the IMLAC display screen were monitored, controlled, and recorded by the PERSEUS system.

The subjects were seated in a darkened room, facing the large cathode-ray tube (CRT) of the IMLAC on which the stimuli were displayed. The distance between the subject's eye and the CRT-face was adjusted so that the 11 X 11 calibration matrix subtended a visual angle of 20 degrees in both vertical and horizontal axes. Each subject was fitted with a metal bite bar surfaced with dental impression compound. The bite bar, firmly attached to the DPIET/display complex, was used to minimize the head-movement and to center subject in the cubic centimeter of space which constitutes the eye-tracker's 'ballpark', i.e., the transitory movement tolerated by the DPIET. The calibration/correction system incorporated in PERSEUS permits the subject, calibrated at the beginning of the session, to leave the bite bar and to return to it with only a re-calibration of the center point. Subjects typed their responses through two keys of a simplified keyboard on which they placed the second and third fingers of their preferred hand; this arrangement eliminated the subject's need to look away from the display--an act which would cause the DPIET to lose track--in order to see the keyboard.

2.3 Experimental Sessions

Each subject had an introductory session or sessions in which he was shown the equipment, and given an explanation of the equipment, procedures, and the purpose of the study. Each subject was then given a trial run as part of his eye-tracker calibration. When the eye-tracker had been adjusted to track the subject's point-of regard throughout the display space, the subject was then instructed in the use of the keyboard through which he was required to enter his manual responses to the exercises. The adults subjects were instructed to look down each column, adding or subtracting as they proceeded, and to try to avoid processing numbers out of the order prescribed by the algorithm.

Subsequent sessions had two parts: calibration of the eye-tracker, followed by the arithmetic exercises. In the calibration phase, a field of 121 dots, in an 11 X 11 array filling the 13 cm X 13 cm display region, was presented. In other words, the calibration rows and columns were two degrees apart. After adjustment of the output voltages of the DPIET, under a program called EYESCAN, the subject was asked to respond to each of the calibration points in sequence by fixating the brightened calibration point and then pressing a key when he was satisfied with his fixation. The eye-pointing data, collected under a program called CALIBRATE, were later used to generate correction filters, via a program called ADAPT, which were used to correct the eye-pointing data collected, via a program called COLLECT, during the subsequent arithmetic session.

The arithmetic exercises were displayed in the standard (row and column) format and projected into the space defined by the 11 X 11 array of calibration points so that each numeral or symbol (except underline) was centered in a cell whose four corners were contiguous calibration points. Underlines were placed above the horizontal midline of cells containing them so that they would be better associated with the digits in the cell above. The digits were approximately 32-mm high by 20-mm wide. A symbol (+,-) for addition or subtraction as appropriate was included in the exercise as presented.

Each new exercise was presented with tentative answer digit placed in the usual location at the base of the first column, i.e., just below the bar line. The tentative answer digit was correct 50 percent of the time; incorrect tentative answer digits were randomly selected to be either one digit higher or one digit lower than the correct answer digit 50 percent of the time. The digit 9 was treated as being 'less' than the digit 0 for this purpose.

The subject was instructed to add the digits in each column and to respond by pressing one key to indicate agreement or an alternative key to indicate disagreement with the machine-proferred answer digit. If he rejected the proferred answer and was correct, the computer responded by changing the answer digit to the correct value. In any case the digit for the next column was then displayed. After the subject had completed the exercise, the computer marked the incorrect digits in the answer row with slashes through the numerals and the correct digits were displayed below them. After a five-second pause, the next exercise was displayed. This procedure was repeated until all 60 exercises had been completed, at, which point the computer informed the subject of the number of correct answers and the number of exercises in the session.

In this study, the subjects were not allowed to delete an answer digit once it had been selected. Also, there was no explicit display of borrowing or carrying information, and no provision to allow the subjects to externally record such information.

2.4 Method for child subjects

The purpose of the study with children was to test the effectiveness of the algorithm used in the initial register- machine model as a predictor of eye-movement performance.

The subjects, two 14-year-old girls and two 11-year-old boys, were selected from a pool of volunteers on the basis of the ability of the DPIET to successfully track their eye movements throughout the display space. This required subjects who could voluntarily limit head movement and whose pupils were relatively large in the testing situation and unobstructed by drooping eyelashes or by drooping upper eyelids.

The children's curriculum consisted of 1000 exercises, of which 600 were addition and 400 subtraction exercises. They were divided into blocks of 100 exercises, composed of 5 randomly generated items from each of 20 specific pre-determined arithmetic exercise structures. Within each block, the order of presentation was randomized; that is, for each position in the set of 100 exercises, a uniform distribution over the remaining exercises was sampled to choose the next exercise, without replacement. The 10 blocks of 100 exercises were then concatenated in a single ordering so that each subject would complete some initial segment of the same ordering.

See Appendix B for a listing of the twenty arithmetic structures used for the children's curriculum.

The addition and subtraction structures were selected to be within a second- to fourth-grade range of difficulty as determined by standard American arithmetic curriculums and studies, which are in close agreement on this topic. Within this range of difficulty, particular structures were selected to represent the range of difficulty ordinarily encountered at this age level (ten to twelve years old).

The equipment and setup for the children were exactly the same as for the adults. The procedures used were similar to those for adults; we note some important differences below.

Calibration of the subject to the eye-tracking system was, in general, more difficult with the younger subjects than with the adults. One problem was that the calibration procedure was for the subjects both time-consuming and attention-demanding. The children were less able to cope with these demands than were the adults. We report below in more detail on studies conducted with the aim of lowering the demands of time and effort placed on subjects during the calibration phase.

Since the largest of the children's exercises did not fill the 13-cm X 13-cm display space, their eye-tracking responses were calibrated using a smaller, relatively centralized, 8-column by 11-row rectangle of points which encompassed the display field in the children's exercises. This matrix of points excluded the more difficult-to-track peripheral calibration points (in particular the corners) of the 11 X 11 array and also significantly reduced the number of calibration points to which the subjects had to respond during the calibration procedure.

In the children's sessions each exercise was first displayed without any proffered answer digits. The subject pressed any key to indicate that he had mentally completed the arithmetic operations on the first column and wanted to see the two possible answer digits for that column. Two digits, selected at the time the curriculum was generated, then appeared on the display. One of the digits was the correct answer and the other was false, chosen on a random basis to be one digit greater or one digit smaller than the correct answer (again with 9 < 0). The proper digits were displayed below the answer cell in the current column, with the smaller digit above the larger digit. By pressing the distal key of the keyboard the subject indicated his choice of the upper proffered digit. Or, he indicated his selection of the bottom proffered digit by pressing the proximal key. In response to his selection, the computer erased the two proffered digits and entered his choice in the answer cell.

Upon completion of each exercise, the computer drew a slash through each incorrect digit and displayed the correct digit below it. A check mark was placed to the left of the answer if the exercise was answered correctly. When the subject was ready to proceed to the next exercise, he tapped any key. The computer responded by erasing the display and presenting a central fixation cross. The subject was instructed to fixate the center of this cross and then to tap any key. The computer then presented the next exercise.

When subjects had completed ten exercises in this fashion, they were allowed to get off the bite bar and rest. Most of the children's sessions were 50 exercises long, though a few sessions were shorter, i.e., 40 or 30 exercises. The number of exercises answered correctly was displayed at the end of the session. No credit was given for partially correct answers.

2.5 Study of Calibration Procedures

A typical problem for eye-tracking devices is that the point of regard as estimated by the eye-tracker may not agree with the subject's own impression of his point of regard. The PERSEUS system is designed to bring the subjective and objective estimates of the point of regard into close agreement. This is accomplished by having the subject look successively at each of 121 calibration points distributed throughout the trackable display space. The CALIBRATE program displays the 11 X 11 array of calibration points in CRT memory mode. It then causes the current calibration target point to be brightened as an indication to the subject that this is the point to be fixated. The subject is instructed to fixate as closely as possible the brightened point and, when he is subjectively satisfied with his fixation, he is to tap a key, which causes the computer to collect 100 samples of DPIET vertical and horizontal output voltages at 2-msec intervals and to locate the fixation point. The coordinates of both the displayed calibration point and the computer-estimated point-of-regard are then stored in a CALIBRATE file and displayed on a separate CRT viewed only by the experimenter. If, in the experimenter's judgment, the subject's fixation response to one of the calibration points appears to be grossly out of line with other local responses, the experimenter can cause the calibration target in question to be brightened again so that the subject can re-fixate it; this type of recalibration of a target point automatically erases the previous record for this point. On the assumption that the normal subject is more likely to be correct about the location of his own fixation, his subjective impression being that he is looking directly at the current fixation point, the target point is taken to be the best estimate of the true point-of-regard and the computer-estimated point-of-regard, if systematically different, is distorted by error. The difference between the two estimates is used to generate corrections for the eye-tracker data collected in the subsequent experimental session. The matrix of calibration points routinely used in experiments with adult subjects consists of 11 rows by 11 columns--with a two-degree separation between adjacent rows and between adjacent columns--and for this reason the calibration procedure is somewhat taxing, particularly for the young subjects. Thus, there is substantial motivation for reducing the total number of calibration points to a minimum and for minimizing the need for recalibrating.

The way that PERSEUS uses the 121 sets of objective and subjective coordinates of the calibration session is to fit the computer-estimated fixation coordinates to the corresponding target display coordinates using a nonlinear regression. The result is one correction surface for horizontal eye-movement components and another correction surface for vertical eye-movement components. These filters are used by the REPORT program to correct the eye-movement data obtained in the related experimental session. PERSEUS has the capacity for computing regressions for models with up to 35 parameters, including powers of terms through the sixth power. Using all 35 parameters, large reductions in the residual error for the correlations of LAG 0 are obtained.

Our concern was to find a simple regression model which would provide adequate predictions of the corrections needed. We assumed initially that, with 121 data points available and with a usual predictively sound ratio of about 1 parameter per 10 data points, a regression model with 12 parameters should be fairly satisfactory. However, our main interest was not in finding the optimal model for prediction, but rather to ensure that the model we used was adequate for the discrimination task called for in the experimental design. Essentially, we were concerned to be able to identify in which 2-degree X 2-degree cell a subject's fixation is contained, out of the 100 cells in the display space.

We decided to test three models of six, nine, and sixteen parameters respectively. By a 6-parameter model, we mean one which uses six parameters to determine the correction for each coordinate. A related 6-parameter model is used to determine the correction for the coordinates in terms of the Euclidean distance measure: $[(x-x')^2 + (y-y')^2]^{1/2}$, which we denote by d. The models for the x coordinate are presented below, with (x,y) indicating the target point and (x',y') indicating the response point.

A. The 6-parameter model

$$x - x' = b_0 + b_1x + b_2y + b_3xy + b_4x^2 + b_5y^2 + e$$

B. The 9-parameter model

$$x - x' = b_0 + b_1x + b_2y + b_3xy + b_4x^2 + b_5y^2 + b_6x^2y + b_7xy^2 + b_8x^2y^2 + e$$

C. The 16-parameter model

$$x - x' = b_0 + b_1x + b_2y + b_3xy + b_4x^2 + b_5y^2 + b_6x^2y + b_7xy^2 + b_8x^2y^2 + b_9x^3 + b_{10}x^3y + b_{11}x^3y^2 + b_{12}x^3y^3 + b_{13}x^2y^3 + b_{14}xy^3 + b_{15}y^3 + e$$

We used the data collected in ten calibration sessions in order to test the three models. The calibration sessions that we considered are labeled in the tables below as: 2D, 2E, 3A, 3B, 3C, 3E, 4A, 4B, 4C, 4D. Each of the ten calibration sessions preceded an experimental session in which this adult subject performed exercises in arithmetic. The numerals refer to the day of the session, and the letters to the sequence of sessions in a particular day. In other words, 3A, 3B, and 3C refer to the first three calibration sessions on the third day for this subject.

These data sets were subjected to three separate analyses. The first two analyses were used to select a model, and the third was used to better gauge the predictive capability of the selected model.

2.5.1 Test of Explanatory Power

The first test was to look at the sum of squared difference between the correction needed to put the recorded point on target and the correction as determined by the model. This difference of differences serves as a measure of the error after correction by the various models. The data obtained for the x and y coordinates and the measure d is contained in Table 2, below, labeled "Sum of Squared Differences -- Lag 0." As one would expect, with increasing number of parameters the sum of squares decreases. However, investigation of the regression coefficients in both the 9- and 16-parameter models, shows that coefficients which have significant t-statistics for the 9-parameter model are not significant in the 16-parameter model. This loss of significance might be an indication that over-fitting is taking place in the 16-parameter model.

It may be seen from Table 2 that the majority of the values lie between 0.8 and 2.4. Values in this range, then, may be taken as a lower bound on the error we can expect when actually predicting coordinates.

Table 2

Sum of Squared Differences -- Lag 0

Coord- \session params	2D	2E	3A	3B	3C	3E	4A	4B	4C	4D
x-6	1.05	1.31	2.48	1.41	1.45	1.31	1.87	1.24	2.32	1.32
x-9	1.03	1.18	2.47	1.32	1.34	1.16	1.80	1.06	2.14	1.19
x-16	0.83	0.90	1.76	0.99	0.99	1.00	1.36	0.84	1.53	0.84
y-6	2.70	2.83	2.23	2.09	2.40	1.47	2.00	1.50	2.19	2.63
y-9	1.99	1.99	1.80	1.62	1.67	0.83	1.43	1.20	1.39	1.74
y-16	1.80	1.74	1.44	1.49	1.59	0.76	1.15	0.98	1.13	1.57
d-6	2.40	2.84	1.56	1.66	1.46	1.71	1.59	1.20	2.74	2.55
d-9	1.83	2.00	1.26	1.43	1.17	1.42	1.57	1.08	1.94	1.67
d-16	1.68	1.61	1.11	1.03	1.04	1.01	1.31	0.65	1.10	0.91

2.5.2 Test of Inter-session Prediction

The second--and more crucial--test was to make a similar comparison of ideal corrections to model-predicted corrections, but across sessions. In this test, the parameters for the model are estimated by regression on the 121 data points from a session and then used to predict corrections for the 121 data points in the next session. In some cases the comparison is across days, but in all cases the subject has gotten off the bite bar between calibration sessions. Since we expect some slight change in the subject's head position in the DPIET to result from the subject's getting off and on the bite bar between calibration sessions, this is a severe test of the predictive abilities of the model. See Table 3 for the sum of squared correction differences, lag 1.

Surprisingly, the 16-parameter model did clearly better than the 9-parameter model at predicting the corrections for subsequent sessions. Out of 27 comparisons between the 9- and 16-parameter models, the 9-parameter model is superior in only 8 cases: prediction for session 3B by regression on 3A for x, y, and d; prediction for 3C by 3D for y and d; prediction for 3E by 3D for x and y; and prediction for 3A by 2D for D. Less surprising is the relative flatness of performance over the range of 6 to 16 parameters.

If the choice of model were to be based strictly on this test, the 16-parameter model would be the obvious choice. However, we felt that the small difference in performance on the test gave us license to consider other factors. One of the factors was the loss of significance of parameters when using the 16-parameter regression. Another factor was the low prior probability assigned to predicting better with a model of more than 12 parameters. Finally, a factor related to the second was

Table 3

Sum of Squared Diffeences -- Lag 1

Coord- \ sessions params	2D-2E	2E-3A	3A-3B	3B-3C	3C-3E	3E-4A	4A-4B	4B-4C	4C-4D
x-6	4.68	5.84	6.21	3.12	2.30	8.58	2.56	5.68	2.21
x-9	4.54	6.01	6.22	2.99	2.12	8.60	2.44	5.60	2.16
x-16	4.45	5.53	6.30	2.71	2.24	8.30	2.38	5.36	2.07
y-6	5.40	9.29	2.91	3.43	1.96	5.57	2.38	4.02	3.94
y-9	4.59	8.91	2.50	2.80	1.40	5.36	2.11	3.21	3.04
y-16	4.43	8.64	2.57	2.83	1.47	5.35	1.95	2.97	2.95
d-6	5.31	4.66	2.94	2.16	2.25	1.92	1.76	5.50	3.18
d-9	4.54	5.00	2.90	1.95	2.04	2.19	1.70	5.13	2.43
d-16	4.31	5.06	2.91	2.15	2.00	2.08	1.52	4.70	2.03

the possiblity of needing to use fewer than 121 data points in calibrating children (as indeed turned out to be the case). For these reasons, we decided to use the 9-parameter model in computing the corrective filters to be used in the scanpath analysis of all the experimental data reported here.

Another important piece of information to be gleaned from Table 3 is a sense of how well we could do if we did not calibrate a subject each time he gets on the bite bar. To give a better sense of the performance of the models as exemplified in Table 3, we provide a rough conversion (see Table 4) from sum-of-squares error to minutes of arc in visual space.

Considering that the cells in which we are trying to place fixations are two degrees square, and that the visual symbols are generally centered in the square, we believe that the error generated by not recalibrating after a subject leaves and then remounts the bite bar is acceptable, though not desirable. It is only rarely that the error could be expected to be more than 20 minutes of arc, which means that when the corrected fixation is in the central half of the cell, we can be virtually certain that the true fixation is in the cell. We made use of this result when it was decided that the child subjects would need to rest off the bite bar periodically during arithmetic sessions.

2.5.3 Test of Intra-session Prediction

After selecting the 9-parameter model on the basis of the strict lag-1 test, we wanted to get some idea of how well we could predict corrections after the calibration part of a session but while the subject was still positioned on the bite bar. This would give us the best sense of the accuracy of the adult data, for which there was a separate complete calibration for each experimental session.

Table 4

Sum of Squares Conversion to Minutes of Arc

Sum of Squares		Minutes of Arc
1.00	-	8.0
2.00	-	10.8
3.00	-	12.4
4.00	-	14.8
5.00	-	17.0
6.00	-	20.0

To make this test, we used the first half of a calibration session (61 points) in a regression to estimate the parameters in the 9-parameter model. We then applied the model with estimated parameters to the last 61 points (one point was used twice for convenience) of the calibration session. The spatial distribution of the two sets of points is roughly equivalent for our purposes.

Table 5 shows both sum of squares (SS) of differences (as in Tables 2 and 3) and mean absolute deviations. The mean absolute deviation (MAD) is the average of the absolute value of the difference between the correction needed and the correction determined by the model. MAD's are much less influenced by the large residuals produced by outliers than the sum of squares statistic. Thus MAD's give a better appraisal of the general fit, while SS's give a better appraisal of how well outliers are fit by the model.

In order to obtain a rough comparison of the results in the third test with the results in the first two tests, we can multiply the SS statistics in Table 5 by two in order to compensate for the difference in size between the data sets. The same procedure can be used to estimate minutes of arc from the SS statistic, using Tables 5 and 4. It will be seen that two times the SS value in Table 5 is greater than 4.0 in only four cases, y and d for sessions 2D and 2E. Thus, by this procedure, we conservatively estimate our within-session accuracy as being approximately 16 minutes of arc.

From a comparison between Tables 5 and 3, we see that intra-session variation accounts for more than half of the correction error. This gives added confidence to the view that we can use a single calibration at the

Table 5

Sum of Squared Differences and Mean Absolute Deviations--Same Day

session	x-Coordinate MAD, SS	y-Coordinate MAD, SS	Euclidean Distance (d) MAD, SS
2D	.046, 0.80	.073, 2.21	.072, 2.06
2E	.060, 1.22	.089, 2.71	.081, 2.21
3A	.051, 1.33	.063, 1.70	.049, 0.82
3B	.049, 0.90	.079, 1.98	.048, 0.86
3C	.052, 0.97	.057, 1.83	.048, 1.24
3E	.051, 0.93	.043, 0.72	.052, 1.01
4A	.048, 1.67	.054, 1.38	.048, 1.35
4B	.039, 0.70	.050, 0.99	.040, 0.66
4C	.051, 1.51	.051, 1.05	.040, 0.90

beginning of a session, even though the subject may dismount and remount the bite-bar during the session.

2.6 Problems Encountered

2.6.1 Problems with Subjects

We had several serious problems in gathering data from young children, and these problems in turn made it difficult to gather as much data as we would have liked from the students of the Ravenswood School District. One of the reasons for selecting the Ravenswood District was its high proportion of minority students. Only one of the subjects for which we have a large number of observations was a minority student from that district.

There are several problems in tracking the movement of young children's eyes with the PERSEUS system. Lids or lashes obscuring the eye are one sort of problem. Dark eyes, which absorb more of the incident light and reflect less are another problem. In general children with obscuring lids had to be excluded, while those with obscuring lashes could sometimes be used if they first curled their lashes out of the way. Dark eyes simply exacerbated any other problems which may have been present, decreasing the likelihood of a successful session.

By far the largest difficulty with very young children is the attempt to use a passive restraint system. Although sounding formidable, the bite bar merely provides the willing subject with a base to help keep steady. Leaning on the bite bar can sufficiently disturb its orientation, and excessive movement of the head around the bite bar, or excessive movement of the body in the chair, are both possible and capable of causing loss of track. Many young children, around the age of eight, seem incapable, even when willing, to remain still enough to even set the system up, let alone keep them in track during a session.

Our chief difficulties with subjects over the age of nine were equipment and software problems. These are described more fully in the section on equipment.

We began with adult subjects from Stanford, and had little trouble gathering data. We then attempted sessions with a very young child (eight) which convinced us to work slowly down beginning with older children. Most of our attempts to gather data from eighth-graders in the Ravenswood district were thwarted by what later turned out to be a problem with the 'fine-tuning' of the SRI eye-tracker. The defect which created small errors in monitoring adult saccades, resulted in loss of track with the children. Only a very small amount of data was gathered in this period.

After the eye-tracker was serviced by SRI, we were able to work more successfully with children. It was during the summer of 1980 that most of the children's data was collected. It was at that time much easier to make arrangement with children from a Palo Alto summer school program, and we tried several volunteers, eventually gathering data from two fifth-graders. We were also fortunate in that one of the eighth-graders from the Ravenswood District agreed to return to allow us to collect substantial data on her eye-movements.

The data collected from individual subjects is massive in character. We collected and analyzed data consisting of more than a quarter of a million eye fixations.

2.6.2 Problems with Computer Systems

There were many problems with the hardware and software for the system we used. The basic task we faced was the merging of an already complicated eye-tracking system with two other computer systems used to present the curriculum. The hardware communication problem was solved by the design and implementation of an digital asynchronous receiver-transmitter (DART), which acted as a communication switch, controlled by the PERSEUS system. In normal mode the IMSSS KI10 computer system could communicate directly with the IMLAC PDS1 system, in order to load programs into the KI10 and PDS1. When experimentation was to start, the PERSEUS system could set the DART to intercept all communication between the KI10 and the PDS1, sending instead to PERSEUS. PERSEUS could then process the communication, and send on appropriate messages to the PDS1 or KI10 as needed.

Software protocols for communicating between the KI10 programs and the PERSEUS programs had to be generated, as well as means for transferring data from the PERSEUS system to the KI10 system for further analysis.

Additional software design and implementation work also went in to overcoming timing problems in the presentation of the curriculum. The KI10 is a timesharing system, and inevitably created delays and inconsistencies in the timing of the presentations. Adjustments were required in the curriculum compilers and interpreters and in the display programs on the PDS1 to minimize the effect of the timesharing environment.

The keyboard for student use also presented a host of problems. We originally expected to use a custom-built keyboard in conjunction with a switch controlled light and a half silvered mirror. The purpose of the proposed arrangement was to let the subject see both the keyboard and the display. With the light off, the subject would see the display through the half-silvered mirror, and when the subject used the knee or foot switch to turn on the light, the keyboard would be reflected in the mirror, and the display would disappear.

A special sized keyboard was designed, built and programmed, and many versions of the above-described system were set up and abandoned. It was finally decided that the half-silvered mirror interfered with the display image ~~too~~ much, and that the light interfered with the eye-tracker. Further, the system was quite complicated. The alternative of having the subject use fewer keys which could be accessed by feel, required changes in the curriculum presentation software.

All of these software and hardware problems caused delays which put off gathering data from any subjects. A further problem caused delays in gathering data from children. That problem was a maladjustment in the SRI DPIET eye-tracker. The main effect of the problem for adults was an over-shooting of the target in saccades. With children, the effect was an increased probability of the tracker losing track. When the DPIET was adjusted, the problem was entirely corrected. Still, additional work was needed to change the calibration routine for children, since tracking children at the periphery of the twenty degree field (where there were no experimental stimulus items, however) remained difficult.

3 Results

We have divided the results into two main sections. The first, deals with the analysis of fixation durations. Here we examine in some detail the probability distribution of fixation durations and also sequential effects of one duration on another, as well as context effects of position in the exercise, etc. In the second section we examine the random-walk for the direction of eye movement described earlier in the theoretical section.

3.1 Analysis of Fixation Durations

The question of whether the fixation durations follow an exponential distribution is of special interest, for this is a consequence of our simple axiom of fixation. This consequence, if true, would have far-reaching implications about the process underlying these durations; as we stated earlier it implies that the process giving rise to these durations is without memory. Also, exponentiality implies independence of the fixation durations on any features of the exercise being performed, such as the point of regard of the fixation, the sum of digits in the column that is being processed, or whether or not a column had a carry from the previous column.

These implications were tested across students. We mention that the exponentiality is easily shown not to hold if we do not include a delay parameter since the fixation durations are necessarily at least 26 msecs long by definition, a number chosen after preliminary analysis of the data and believed to be consistent with other eye-movement data and analyses. The data are not highly sensitive to the exact number chosen.

The process of interest is then thought of as having two components, a fixed waiting time, followed by a process with exponential waiting times. Therefore the delay parameter has to be estimated. Since we are typically dealing with large data sets of many thousand points (typically about 4,000), it is certainly sufficient to estimate this parameter for n data points by taking the minimum fixation duration and multiplying it by n/(n+1).

An immediate but weak test for exponentiality is the closeness of the mean to the standard deviation which are equal for the exponential distribution (but are also equal for many other distributions). We have the following table of means and standard deviations in seconds. (Remember the minimum multiplied by n/(n+1) has been subtracted out.) In this table, KJ and JF refer to the adult subjects, JM and CJ to the eighth-grade girls, and CH and JU to the fifth-grade boys. The numeral following the two letters refers to the number of the session, numbered separately for the addition and subtraction sessions--referred to by A and S, respectively, after the numeral--for the two adult subjects.

It appears that the standard deviation is almost always smaller than the mean--there are only five contrary cases. This indicates at least two possibilities. One possibility is that the tail of the distribution is thinner than that of the exponential. The other possibility is that the mass which for the exponential distribution would be near zero has been shifted to the right. We shall return to this later.

Table 6

Estimated Means and Standard Deviations of Fixation Duration

for Individual Subject Sessions

	m	s		m	s
KJ1A	.3378	.3204			
KJ2A	.3338	.3142	JF2A	.3020	.2460
KJ3A	.2918	.2652	JF3A	.2796	.2300
KJ4A	.2168	.2146	JF4A	.2835	.2227
KJ5A	.2834	.2659	JF5A	.2615	.2262
KJ6A	.2553	.2261	JF6A	.2721	.2232
KJ7A	.2316	~~.2032~~	JF7A	.2293	.2219
KJ8A	.2379	.2048	JF8A	.2675	.2248
KJ9A	.3210	.2933	JF9A	.2694	.2212
KJ10A	.3232	.3014	JF10A	.2213	.1972
KJ11A	.2784	.2586	JF11A	.3027	.2730
KJ12A	.2711	.2629	JF12A	.3008	.2479
			JF1S	.2795	.2094
KJ2S	.2527	.2268	JF2S	.2768	.1990
KJ3S	.2368	.2053	JF3S	.2807	.2183
KJ4S	.2668	.2227	JF4S	.2985	.2466
KJ5S	.3237	.2838	JF5S	.2873	.2214
KJ6S	.2582	.2440	JF6S	.2829	.2353
KJ7S	.1293	.1336	JF7S	.2903	.2432
KJ8S	.1813	.1689	JF8S	.2982	.2386
KJ9S	.2425	.2215	JF9S	.2713	.2438
KJ10S	.2379	.2300	JF10S	.2436	.2158

Table 6, continued

JM1B	.2048	.1802	CH1B	.1572	.1450
CJ2B	.2714	.1978	CH2B	.1422	.1380
CJ3B	.2433	.1639	CH3B	.1428	.1421
CJ4B	.2390	.1602	CH4B	.1444	.1461
CJ5B	.2552	.1746	CH5B	.1493	.1412
CJ6B	.2601	.1761	CH6B	.1408	.1324
CJ7B	.2541	.1695	CH7B	.1424	.1420
CJ8B	.2429	.1800			
CJ9B	.2671	.2025			
CJ10B	.2487	.1726			
CJ11B	.2424	.1794			
CJ12B	.2495	.1764			
CJ13B	.2512	.1736			
JU01B	.1661	.1596	JU06B	.1528	.1490
JU02B	.1516	.1531	JU07B	.1611	.1670
JU03B	.1571	.1559	JU08B	.1473	.1603
JU04B	.1746	.1625	JU09B	.1868	.1854
JU05B	.1597	.1592	JU10B	.1843	.1921

3.1.1 Correlation of Successive Fixation Durations

The second test concerned the correlation of successive fixation durations. We show in Table 7 the auto-correlations of lag one for full sessions for the adult subjects. The statistic can be written as follows where d_i represents the ith fixation duration, and $\bar{d}$ their average.

$$\sum_{i=2}^{n} (d_i - \bar{d})(d_{i-1} - \bar{d}) \Big/ \mathrm{sqrt}\Big[\Big(\sum_{i=1}^{n} (d_i - \bar{d})^2\Big) \Big(\sum_{i=2}^{n} (d_{i-1} - \bar{d})^2 \Big)\Big]$$

The second row of data in Table 7 shows the actual number of fixation durations in each session. As is evident, there is considerable variation in the number, especially across subjects.

For the data of KJ it seems that there is a very small positive correlation since 15 out of the 17 sessions have a positive correlation. The data for JF seem to have a small negative correlation, if any at all. Overall we may conclude that there is really no significant effect of the length of a fixation on that of the successive fixation. This finding is therefore consistent with the (delayed) exponential model.

3.1.2 Chi-square Test of Exponential Distribution

To further evaluate the exponential model we decided to use a more global test. In standard fashion, we divided up the range of the distribution into n intervals, determined the number of data points that fall in each interval, denoted n_i, where i denotes the ith interval, and determined the expected number of data points in each interval, denoted exp_i. The intervals were chosen in order to have the exp_i nearly constant. The statistic $\sum(n_i - exp_i)^2 / (exp_i)$ has an asymptotic chi-square distribution with n-2 degrees of freedom, because the final cell is determined from the n-1 previous cells, and we must estimate the parameter of the exponential distribution. We tried this test on the addition and subtraction files, with results at least as pessimistic as those tabulated below, on nine degrees of freedom.

JF2A	JF3A	JF4A	JF5A	JF6A	JF7A	JF8A
1127	1316	861	848	998	1119	1080

From Table 6 and these results it was clear that the simple delayed exponential model was not going to fit well.

3.1.3 Alternative Models

Elimination of possible contaminants. Four methods of altering the models were attempted. First, it is possible that the data are contaminated with nonfixations or saccades. Assuming that these contaminants have an exponential distribution we can estimate a mixing parameter, which indicates the degree of contamination, and two exponential parameters. This model was not attempted but was replaced with the following model which was computationally easier to implement. If the mass of the distribution of the contaminants is probabilistically nearly disjoint from that of the fixation durations we could then guess a cutoff point and examine only data which were greater than that point.

Random delay. Second, it is possible that there is a process which causes a delay that is not fixed but random. Assuming that the distribution of the delay is also exponential we need to estimate two parameters of a distribution which is the convolution of two exponential distributions with different parameters.

Table 7

Estimated Correlations of Successive Fixation Durations

for Individual Subject Sessions

	JF2A	JF3A	JF4A	JF5A	JF6A	JF7A	JF8A	JF9A	JF10A	JF11A	JF12A
correl	.032	-.046	-.060	-.034	-.029	.048	.005	-.014	.040	.037	.0146
count	4723	4698	3114	3933	4359	4629	4454	4790	5789	4523	4954

	KJ1A	KJ2A	KJ3A	KJ4A	KJ5A	KJ6A	KJ7A	KJ8A	KJ9A	KJ10A	KJ11A	KJ12A
correl	.022	.056	.027	.071	.031	.085	.016	.029	-.032	.011	.034	.041
count	7399	5583	6406	8194	6304	6762	6616	6916	4624	6108	6206	6521

	JF1S	JF2S	JF3S	JF4S	JF5S	JF6S	JF7S	JF8S	JF9S	JF10S
correl	-.022	-.050	-.018	-.050	-.049	-.017	.014	-.036	-.005	.040
count	2686	2292	2317	2574	2467	2353	2322	2218	2435	2312

	KJ2S	KJ3S	KJ4S	KJ5S	KJ6S	KJ7S	KJ8S	KJ9S	KJ10S
correl	.027	.008	-.037	-.053	.018	.156	.035	-.024	.032
count	3172	3362	2917	2678	2918	4842	3814	3146	3494

	JM1B	CH1B	CH2B	CH3B	CH4B	CH5B	CH6B	CH7B
correl	.090	.120	.109	.125	.139	.077	.114	.132
count	2561	5146	4163	3297	3482	3799	4981	3658

	CJ2B	CJ3B	CJ4B	CJ5B	CJ6B	CJ7B	CJ8B	CJ9B	CJ10B
correl	.060	.070	.059	.094	.059	.066	.041	-.000	.022
count	1383	1605	1564	1221	1381	1206	1361	1395	1261

	CJ11B	CJ12B	CJ13B
correl	.018	.031	.047
count	1095	1196	1045

	JU01B	JU02B	JU03B	JU04B	JU05B	JU06B	JU07B	JU08B	JU09B	JU10B
correl	.043	.086	.081	.094	.098	.096	.079	.101	.096	.097
count	3427	3228	2883	2894	3091	2710	3149	3420	3494	2890

Gamma distribution. Third, it is possible that the fixations are the result of a convolution of exponentials with the same parameter. This gives rise to the gamma distribution. These last two models both have interesting metrics on the distance from exponentiality. The convolution model has the parameter for the expectation of the first process. If this is, say, 10 msecs, we see that except for a process of a very short duration we have nearly an exponential process. In the case of the gamma distribution we estimate two parameters, γ and ρ. The parameter ρ has the interpretation of the number of exponentials that have been convolved in order to form the process. The closer the estimated ρ is to one, the closer the process is to an exponential process. This is because the form for the gamma distribution includes the exponential as a special case, when ρ is one. We mention here that we used maximum-likelihood estimates which were found by a gradient search routine that started at the method-of-moments estimate.

In Table 8 we show the results of the model with a cutoff point. The cutoff point was decided to be .040 seconds on the basis of examining histograms of the distribution of fixation durations between 0 and .100 seconds. From these, a cluster was apparent which was consistently reduced in size by 40 msecs. The first column of data in Table 8 shows the ratio of the reduced sample size to the original sample size when the cutoff is imposed. The second column shows the chi-square value for 19 degrees of freedom.

Next we show in Table 9 the results for the convolution of two exponentials. The method of maximum-likelihood was used to estimate the two parameters of the convolution, λ_1 and λ_2. Here the parameters are the reciprocals of the means of the two distributions. Finally we show in

Table 8

Chi-square Test of the Exponential Cutoff Model

with 19 Degrees of Freedom

	Ratio of sample size	χ^2		Ratio of sample size	χ^2
			JF1S	2540/2686	412.9
JF2A	4383/4690	578.2	JF2S	2197/2292	362.9
JF3A	4414/4698	731.0	JF3S	2219/2317	364.8
JF4A	2961/3114	526.9	JF4S	2445/2574	369.7
JF5A	3609/3933	403.0	JF5S	2364/2467	366.7
JF6A	4055/4359	646.9	JF6S	2201/2353	291.5
JF7A	4015/4629	228.6	JF7S	2177/2322	302.0
JF8A	4118/4454	559.9	JF8S	2113/2218	353.1
JF9A	4470/4790	700.0	JF9S	2224/2435	189.2
JF10A	5026/5789	422.9	JF10S	2041/2312	135.8
JF11A	4188/4523	1284.0			
JF12A	4707/4954	1262.0			
KJ1A	[illegible]730/7399	191.2			
KJ2A	5099/5583	134.8	KJ2S	2862/3172	129.0
KJ3A	5782/6406	220.9	KJ3S	3011/3362	164.9
KJ4A	6817/8194	99.5	KJ4S	2672/2917	164.1
KJ5A	5666/6304	138.7	KJ5S	2547/2678	261.8
KJ6A	6012/6762	245.4	KJ6S	2601/2918	77.5
KJ7A	5786/6616	198.0	KJ7S	3532/4842	25.3
KJ8A	6051/6916	233.7	KJ8S	3134/3814	96.77

Table 8, continued

KJ9A	4244/4624	194.6	KJ9S	2796/3146	153.9
KJ10A	5609/6108	173.8	KJ10S	3068/3494	157.9
KJ11A	5372/6206	136.4			
KJ12A	5797/6521	14[illegible]8			
JM1B	2202/2561	163.3	CH1B	4000/5146	316.7
CJ2B	1268/1383	365.2	CH2B	3038/4163	130.9
CJ3B	1474/1605	396.1	CH3B	2365/3297	119.4
CJ4B	1436/1564	386.7	CH4B	2563/3482	127.2
CJ5B	1151/1221	285.0	CH5B	2856/3799	123.1
CJ6B	1307/1381	293.6	CH6B	3564/4981	124.7
CJ7B	1128/1206	287.8	CH7B	2663/3658	65.9
CJ8B	1258/1361	258.6			
CJ9B	1301/1395	252.9			
CJ10B	1175/1261	292.7			
CJ11B	999/1095	235.3			
CJ12B	1123/1196	285.1			
CJ13B	952/1045	306.8			
JU01B	2670/3427	82.5	JU06B	2024/2710	40.9
JU02B	2442/3228	42.0	JU07B	2368/3149	42.6
JU03B	2170/288[illegible]	48.8	JU08B	2498/3420	37.3
JU04B	2305/2894	50.8	JU09B	2[illegible]19/3494	34.1
JU05B	2383/3091	44.4	JU10B	2251/2890	38.3

Table 9

Chi-square Test of the Convolution of Two Exponentials with 19 Degrees of Freedom

	λ_1 ($\times 10^6$)	λ_2	χ^2		λ_1 ($\times 10^6$)	λ_2	χ^2
				JF1S	0.950	3.612	592.0
JF2A	1.796	3.300	758.7	JF2S	0.853	3.612	652.2
JF3A	1.796	3.576	997.7	JF3S	0.698	3.576	557.5
JF4A	0.948	3.540	761.1	JF4S	0.853	3.333	525.0
JF5A	1.551	3.837	541.8	JF5S	0.810	3.470	578.5
JF6A	1.796	3.686	866.1	JF6S	0.774	3.540	397.7
JF7A	1.796	4.371	270.8	JF7S	0.857	3.435	417.3
JF8A	1.710	3.723	748.0	JF8S	0.810	3.367	488.2
JF9A	1.886	3.723	967.4	JF9S	0.900	3.686	239.2
JF10A	2.183	4.503	480.9	JF10S	0.857	4.116	170.6
JF11A	1.474	3.300	1312.0				
JF12A	1.886	3.333	1614.0				
KJ1A	2.925	2.954	228.6				
KJ2A	2.407	2.984	172.8	KJ2S	1.216	3.954	191.0
KJ3A	2.4[illegible]7	3.435	269.9	KJ3S	1.340	4.242	224.4
KJ4A	3.556	4.593	112.1	KJ4S	1.047	3.760	250.4
KJ5A	2.527	3.540	195.4	KJ5S	0.990	3.076	400.5
KJ6A	2.925	3.915	307.1	KJ6S	1.158	3.876	112.6
KJ7A	2.527	4.327	260.0	KJ7S	2.401	7.706	58.8
KJ8A	2.786	4.200	292.3	KJ8S	1.544	5.494	106.0

Table 9, continued

KJ9A	1.886	3.107	256.1	KJ9S	1.158	4.116	198.4
KJ10A	2.292	3.107	230.6	KJ10S	1.407	4.200	190.8
KJ11A	2.653	3.576	191.8				
KJ12A	2.786	3.686	181.7				
JM1B	1.050	4.876	186.7	CH1B	2.292	6.379	361.4
CJ2B	0.463	3.686	419.1	CH2B	1.698	7.046	199.5
CJ3B	0.569	4.116	491.4	CH3B	1.397	6.977	204.7
CJ4B	0.569	4.200	674.6	CH4B	1.463	6.907	154.0
CJ5B	0.359	3.915	392.7	CH5B	1.524	6.704	181.9
CJ6B	0.418	3.845	407.6	CH6B	2.281	7.117	238.8
CJ7B	0.418	3.935	395.4	CH7B	1.524	7.046	144.0
CJ8B	0.463	4.116	349.3				
CJ9B	0.488	3.744	346.4				
CJ10B	0.440	4.021	383.9				
CJ11B	0.397	4.126	289.2				
CJ12B	0.397	4.008	386.3				
CJ13B	0.341	3.994	338.7				
JU01B	1.397	6.009	103.9	JU06B	1.050	6.545	53.2
JU02B	1.333	6.572	79.4	JU07B	1.216	6.209	82.4
JU03B	1.103	6.365	87.2	JU08B	1.407	6.790	77.6
JU04B	1.209	5.718	80.6	JU09B	1.407	5.352	44.8
JU05B	1.276	6.260	64.5	JU10B	1.103	5.426	71.2

Table 10 the results for the gamma model, with again the two parameters γ and ρ estimated by maximum likelihood. Evidently the gamma distribution fits the data much better than the convolution of two exponentials.

Inclusion of saccade time. Fourth, a question arose as to whether or not the length of the fixation should include the temporal length of the saccade immediately preceding it. It was thought that possibly the nonexponentiality was due to the exclusion of that time. However, when the time of the previous saccade was adjoined in the analysis of a few of the sessions, the only effect was a further departure from exponentiality as seen in the histograms.

3.1.4 Effects of Structural Features

Given the departure from an exponential distribution of fixation durations, it is interesting to investigate further how to model the ways the distribution departs from the exponential. One way to accomplish this is to determine the effect of certain structural features, if any, on the distribution of the fixation durations. To do this we regressed the fixation duration on the following variables:

1) ROW - the row the fixation is in,
2) COL - the column the fixation is in,
3) LENGTH - the number of digits in the top row,
4) ONOFF - an indicator in subtraction of whether a borrow was needed from the next column, and in addition of whether a carry was given to that column from the previous one.

Table 10

Chi-square Test of the Gamma Distribution

with 19 Degrees of Freedom

File	γ	ρ	χ^2	File	γ	ρ	χ^2
				JF1S	5.976	1.662	183.8
JF2A	4.491	1.347	343.6	JF2S	6.290	1.735	132.2
JF3A	5.285	1.478	404.7	JF3S	6.089	1.723	148.7
JF4A	5.918	1.691	222.6	JF4S	5.209	1.566	190.1
JF5A	4.913	1.277	267.7	JF5S	5.862	1.684	151.6
JF6A	5.463	1.476	367.3	JF6S	5.109	1.455	141.0
JF7A	4.356	1.008	264.2	JF7S	4.807	1.385	194.1
JF8A	5.296	1.407	559.9	JF8S	5.237	1.561	165.6
JF9A	5.506	1.473	394.3	JF9S	4.566	1.239	131.8
JF10A	4.794	1.060	402.2	JF10S	4.631	1.124	121.5
JF11A	4.562	1.389	857.0				
JF12A	5.197	1.573	709.5				
KJ1A	3.189	1.061	179.1				
KJ2A	3.280	1.078	117.2	KJ2S	4.913	1.191	89.8
KJ3A	3.750	1.081	188.1	KJ3S	4.917	1.160	116.6
KJ4A	4.706	1.010	113.5	KJ4S	4.881	1.356	119.8
KJ5A	3.808	1.066	137.7	KJ5S	4.118	1.341	210.5
KJ6A	4.296	1.086	209.8	KJ6S	4.238	1.090	70.9
KJ7A	4.908	1.129	134.4	KJ7S	7.741	1.000	59.8
KJ8A	4.473	1.060	222.6	KJ8S	5.556	1.002	105.7
KJ9A	3.631	1.158	143.1	KJ9S	4.743	1.139	128.1

Table 10, continued

KJ10A	3.557	1.140	115.4	KJ10S	4.498	1.060	162.4
KJ11A	3.862	1.069	132.2				
KJ12A	3.821	1.023	161.6				
JM1B	5.009	1.022	174.6	CH1B	6.380	1.006	359.6
CJ2B	5.536	1.492	246.7	CH2B	7.070	1.002	201.4
CJ3B	6.257	1.524	242.4	CH3B	7.074	1.010	119.4
CJ4B	6.217	1.487	266.5	CH4B	6.965	1.000	154.9
CJ5B	7.271	1.846	129.6	CH5B	6.790	1.008	184.9
CJ6B	6.700	1.750	141.9	CH6B	7.086	1.000	240.4
CJ7B	6.300	1.600	178.5	CH7B	7.067	1.006	148.7
CJ8B	6.200	1.500	164.4				
CJ9B	5.700	1.510	159.5				
CJ10B	6.000	1.500	178.3				
CJ11B	5.900	1.370	169.6				
CJ12B	6.500	1.630	158.2				
CJ13B	6.237	1.555	181.7				
JU01B	6.019	1.003	104.3	JU06B	6.586	1.002	54.1
JU02B	6.570	1.000	80.0	JU07B	6.177	1.000	82.9
JU03B	6.368	1.006	90.2	JU08B	6.832	1.000	79.0
JU04B	5.716	1.005	81.0	JU09B	5.337	1.006	45.2
JU05B	6.312	1.004	66.0	JU10B	5.394	1.000	71.5

The regression equation therefore was

$$y_i = a_0 + a_1 \mathrm{LENGTH}_i + a_2 \mathrm{ROW}_i + a_3 \mathrm{COLUMN}_i + a_4 \mathrm{ONOFF}_i$$

where y_i denotes the fixation duration of the ith data point.

The results of these regressions are shown in Table 11. An asterisk denotes a nonsignificant coefficient at the 97.5% level (one-sided t-test).

We can see the following points from an inspection of Table 11. First, as the number of the row the subject is in increases, the fixation duration also increases, for the coefficient is almost uniformly positive. Since the correlation of the variables involved is small we may be inclined to surmise that the subject lingers more further down the column. This may have to do with the fact that the numbers added are now much larger. This effect is stronger in addition exercises than in subtraction exercises, which is not surprising since there are only two rows of digits in subtraction. Column and length are often nonsignificant and change sign with subject. Little more of a concrete nature can be determined. Finally, onoff for subtraction, which is borrowing, has a clear positive effect on fixation duration, which might be guessed considering that borrowing may be the most difficult algorithmic step in subtraction (indeed in all the exercises we presented to subjects).

As would be expected from the size of the various coefficients, the regression on the structural variables accounts for only a relatively small part of the variation in the fixation durations--the range of the square of the multiple correlation (R^2) is .000324 to .09123. These results stand in contrast to the regression of response and latency data on the same structura variables where R^2s of .70 or greater are common (for extensive analyses, see Suppes & Morningstar, 1972).

Table 11

Regression Coefficients of Some Structural

Variables for Fixation Durations

	JF2A	JF3A	JF4A	JF5A	JF6A	JF7A	JF8A	JF9A	JF10A	JF11A	JF12A
row	.031	.034	.030	.037	.033	.022	.031	.024	.014	.042	.028
column	.019	.006	*	*	*	*	*	*	*	*	.009
length	-.018	-.016	-.011	-.016	-.010	-.037	-.022	-.012	-.019	-.031	-.012
onoff	-.037	-.033	*	*	*	*	*	*	*	*	*

	KJ1A	KJ2A	KJ3A	KJ4A	KJ5A	KJ6A	KJ7A	KJ8A	KJ9A	KJ10A	KJ11A	KJ12A
row	.025	.015	.015	.005	.017	.012	.015	.017	.027	.009	.021	.019
column	-.022	-.020	-.021	-.028	-.021	-.008	-.015	-.012	-.032	-.015	-.013	-.019
length	.012	.011	.019	.017	.015	.030	.011	*	.017	.008	.008	-.030
onoff	*	.030	.025	.020	*	.026	.030	.035	.042	.056	.029	.023

	JF1S	JF2S	JF3S	JF4S	JF5S	JF6S	JF7S	JF8S	JF9S	JF10S
row	.032	.033	.025	.046	.035	.034	.039	.033	.033	.016
column	.012	*	.020	.013	.015	*	.016	*	*	.013
length	-.023	*	-.028	-.009	*	*	*	-.017	*	-.030
onoff	.031	.028	.025	.044	*	*	*	*	.026	.026

	KJ2S	KJ3S	KJ4S	KJ5S	KJ6S	KJ7S	KJ8S	KJ9S	KJ10S
row	.022	.023	.018	.011	.009	.013	.016	.019	.014
column	*	*	*	-.015	-.025	-.008	*	-.013	*
length	*	*	*	*	.015	.013	*	*	.025
onoff	.042	.047	.057	.058	.064	.016	.020	.034	.055

Table 11, continued

	JM1B	CH1B	CH2B	CH3B	CH4B	CH5B	CH6B	CH7B	CJ2B	CJ3B
row	.007	.006	*	-.006	*	-.003	-.005	-.003	*	*
column	*	-.007	-.006	*	.005	*	-.007	-.008	*	*
length	*	*	*	*	*	*	-.014	*	-.020	*
onoff	*	*	.015	.011	.023	*	*	.013	*	*

	CJ4B	CJ5B	CJ6B	CJ7B	CJ8B	CJ9B	CJ10B	CJ11B	CJ12B	CJ13B
row	*	-.007	*	*	*	*	*	*	*	-.011
column	*	*	*	.015	.012	*	*	*	*	.018
length	*	*	*	-.022	-.020	*	*	*	*	-.033
onoff	*	.048	.027	.027	.045	.029	*	*	*	*

	JU01B	JU02B	JU03B	JU04B	JU05B	JU06B	JU07B	JU08B	JU09B	JU10B
row	.012	.014	.014	.017	.016	.013	.015	.020	.024	.023
column	*	*	*	*	-.008	*	-.008	-.007	-.011	-.011
length	.010	*	*	.018	*	*	.022	.012	.018	*
onoff	*	*	*	.019	.024	*	*	*	*	.023

3.1.5 Mixture Distribution

After investigation of the many different aspects of the distribution of fixation durations described above, for reasons given in the theoretical discussion we fitted the mixture distribution with three parameters. The results of this analysis are shown in Table 12. As previously, the results are given for individual sessions. The first column shows the estimated parameter of the exponential distribution and the second, the estimated parameter of the convolution of two exponential distributions. The third column shows the mixture weighting α on the exponential distribution. It should be noted that if the convolution of two exponential distributions with the parameter λ_2 is the same exponential distribution as estimated in column 2, then the coefficient for λ_2 should be one-half that for λ_1. In the fourth column we show the chi-square fit of the mixture distribution, which has one more parameter than the gamma distribution, and we might expect somewhat better fit. It is evident from comparison of Table 10 and Table 12 that the mixture distribution actually fits very much better than the gamma distribution for almost all sessions, and for only one of the 72 sessions analyzed is the gamma distribution actually better (session JU09B), and then only very slightly.

In the fifth column of Table 12 we give the estimate of the average error statistic for the mixture model. In this case the average error is the square root of the sum over cells of the squared difference between the true and predicted probability of an arbitrary fixation duration being within the cell, divided by the predicted probability. The maximum-likelihood estimate of the average error is computed from the Chi-square statistic by the following equation, where χ^2 is the Chi-square statistic,

Table 12

Chi-square and Average Error Statistics for the Mixture of Exponential and Convolution of Exponentials Model with 18 Degrees of Freedom

	λ_1	λ_2	α	χ^2	Average error
JF2A	.302	.151	.33	194.1	.194
JF3A	.330	.130	.27	269.3	.232
JF4A	.343	.137	.17	167.3	.220
JF5A	.277	.126	.38	141.8	.178
JF6A	.297	.131	.29	257.0	.235
JF7A	.229	.115	.66	124.6	.153
JF8A	.288	.129	.32	187.9	.196
JF9A	.294	.130	.29	270.7	.230
JF10A	.211	.116	.57	187.7	.172
JF11A	.368	.136	.33	675.7	.382
JF12A	.386	.140	.18	535.9	.324
JF1S	.275	.140	.19	136.2	.211
JF2S	.256	.143	.12	57.6	.135
JF3S	.351	.135	.14	111.4	.203
JF4S	.393	.134	.22	123.8	.204
JF5S	.332	.139	.17	111.9	.197
JF6S	.308	.136	.28	88.7	.176
JF7S	.330	.135	.32	122.2	.213

Table 12, continued

JF8S	.343	.144	.22	114.4	.210
JF9S	.296	.126	.48	80.0	.162
JF10S	.179	.152	.48	83.3	.170
KJ1A	.353	.149	.74	117.5	.117
KJ2A	.349	.147	.71	65.8	.095
KJ3A	.292	.146	.66	98.3	.113
KJ4A	.192	.188	.86	106.8	.105
KJ5A	.293	.132	.69	61.3	.085
KJ6A	.235	.143	.61	90.7	.105
KJ7A	.212	.131	.61	46.4	.068
KJ8A	.203	.144	.58	73.1	.091
KJ9A	.336	.151	.59	83.6	.121
KJ10A	.338	.147	.64	64.0	.089
KJ11A	.288	.129	.70	66.8	.091
KJ12A	.286	.111	.75	80.3	.099
KJ2S	.253	.126	.55	55.6	.112
KJ3S	.237	.118	.53	57.8	.112
KJ4S	.242	.143	.45	91.0	.160
KJ5S	.389	.142	.37	143.5	.218
KJ6S	.268	.119	.66	34.3	.082
KJ7S	.074	.110	.62	58.9	.094
KJ8S	.126	.137	.62	74.5	.124
KJ9S	.243	.121	.62	88.2	.151
KJ10S	.258	.100	.68	104.6	.159

Table 12, continued

JM1B	.185	.117	.60	84.7	.164
CJ2B	.200	.145	.24	196.7	.361
CJ3B	.158	.132	.19	174.1	.313
CJ4B	.155	.130	.19	206.2	.348
CJ5B	.165	.133	.11	108.6	.275
CJ6B	.175	.135	.13	111.6	.263
CJ7B	.149	.137	.14	134.1	.312
CJ8B	.203	.126	.24	115.6	.270
CJ9B	.237	.139	.25	107.8	.256
CJ10B	.189	.129	.19	120.9	.288
CJ11B	.192	.131	.28	113.7	.298
CJ12B	.205	.130	.17	121.8	.297
CJ13B	.166	.136	.20	137.3	.340
CH01B	.112	.114	.61	291.9	.231
CH02B	.087	.116	.61	190.3	.204
CH03B	.018	.096	.27	49.8	.102
CH04B	.089	.117	.62	152.5	.198
CH05B	.089	.115	.57	172.2	.202
CH06B	.081	.120	.62	223.2	.204
CH07B	.023	.096	.28	54.2	.103
JU01B	.111	.128	.62	91.7	.148
JU02B	.092	.126	.62	78.4	.139
JU03B	.093	.134	.62	81.9	.151

Table 12, continued

JU04B	.120	.132	.62	69.6	.136
JU05B	.100	.130	.63	63.5	.124
JU06B	.093	.126	.62	48.7	.111
JU07B	.091	.141	.63	71.8	.133
JU08B	.077	.139	.64	34.8	.076
JU09B	.122	.148	.63	45.7	.093
JU10B	.109	.162	.65	61.6	.126

k the degrees of freedom, and N the sample size:

$$e = [(\chi^2 - \frac{(k-1)^2}{2})/N\chi^2]^{1/2}$$

The average error is a good measure of how well a model is fitting, when a Chi-square test would reject the model, because the Chi-square test sets ever more stringent criteria on the average discrepancy as the sample size goes up (see, Kraemer, 1965).

We see in Table 12 that the fit as indicated by average error is not too bad. Although the adult average error ranges between 7 and 38 percent, in 30 out of 42 cases it is below 20 percent. The children's fit is not as good, with a range of 8 to 36 percent, and only 14 out of 30 sessions under 20 percent. Still, almost all the problems of fit for the children are with subject CJ, for whom the mixture model is clearly less adequate than it is for all the other subjects.

3.2 Random-walk Model of Movement Direction

3.2.1 Fit to register-machine model.

The next process of interest was that of the sequential grid positions attended to by the subjects, i.e., the grid scanpath. In order to determine the degree of fit of the scanpaths to the register-machine model it was necessary to create some measures of goodness-of-fit. Since there seem to be no close precedents to help guide us, inevitably the measures that we developed are to some extent *ad hoc*.

The first measure is denoted corr1, and is constructed as follows. The theoretical scanpath for the register-machine model was simplified for this analysis. A perfect scan-path in terms of this model was defined to be moving top-to-bottom in each column with the fixation in the bottom of a column being immediately succeeded by a fixation in the first row of the column immediately to the left. This sequence begins in the upper right grid position. A four-by-four exercise is shown below with its proper scanpath to the right.

3	6	7	7	*	(25)	(19)	(13)	(7)	(1)
2	9	0	3	*	(26)	(20)	(14)	(8)	(2)
3	5	5	1	*	(27)	(21)	(15)	(9)	(3)
4	4	0	8	*	(28)	(22)	(16)	(10)	(4)
-	-	-	-	*	(29)	(23)	(17)	(11)	(5)
? ?	?	?	?	*	(30)	(24)	(18)	(12)	(6)

Similar correspondences can of course be constructed for exercises of other configurations. In accordance with the register-machine model for addition we denote the duration of the fixation at time t as "good" when the position p_t of a fixation at time t is either p_{t-1} or $p_{t-1} + 1$, i.e., staying put in the same grid, p_{t-1}, or advancing one square, $p_{t-1} + 1$; otherwise we denote the time "bad". (The analysis is slightly different for subtraction.) The statistic is simply the ratio of the "good" time to the overall time. This measure is in a sense optimistic since a large fraction of the fixations are at the same grid position as the previous fixation, partly due to the tight fixation parameters used in the data reduction.

Because we would not claim that the scan paths nearly agreed with those predicted for the register-machine model if 90% of the fixation positions were repeats and the other 10% appeared to be independent of the model, we constructed corr2. The only difference between corr1 and corr2 is that repeat time is not counted as "good" or "bad" time in the determination of corr2. For this reason corr2 is pessimistic in that some repeating should be included in "good" time.

Corr3 is a third measure of goodness-of-fit for scan paths. To compute this measure we first determine how many fixation positions, say n, occurred for a particular exercise. Then two sequences are created. The sequence 1,...,n and the sequence of grid positions predicted by the register-machine model, as indicated above. Then the correlation of these two sequences is calculated. This is corr3. The statistic is then averaged across the exercises for an entire session. Corr3 is probably slightly pessimistic in that it should really be calculated as the maximum correlation between the sequence of grid positions and any monotonic transformation of the sequence {1,...,n}, since this would naturally allow for repetition.

We point out that these measures all increased with the length of the exercise, most notably corr2 on addition sessions. Our analysis allowed for the underline position to be analyzed as a separate location or joined to the answer position immediately below. The analysis shown in Table 13 corresponds to the separation of the underline but the alternative seems reasonable as well.

In the case of addition, the average of corr1 and corr2 seems to be a reasonable estimate of goodness-of-fit. The fit as measured by corr1 for the two adult subjects is quite good. For subtraction the measures are so disparate it is difficult to say anything with assurance. The fit as measured by corr3 is surprisingly good for almost all 72 sessions, but we believe that it is not model-specific enough for the main theoretical framework of our analysis.

Table 13

Three Measures of Fit of the Scanpaths
to the Register-machine Model

	KJ1A	KJ2A	KJ3A	KJ4A	KJ5A	KJ6A	KJ7A	KJ8A	KJ9A	KJ10A	KJ11A	KJ12A
corr1	.872	.871	.891	.793	.892	.904	.899	.894	.858	.879	.890	.896
corr2	.594	.619	.632	.388	.633	.650	.656	.670	.658	.633	.653	.645
corr3	.908	.862	.864	.843	.828	.853	.785	.798	.791	.829	.825	.814

	JF2A	JF3A	JF4A	JF5A	JF6A	JF7A	JF8A	JF9A	JF10A	JF11A	JF12A
corr1	.787	.817	.791	.812	.841	.810	.818	.829	.841	.841	.824
corr2	.495	.510	.475	.511	.562	.494	.506	.535	.545	.552	.518
corr3	.857	.857	.869	.839	.903	.852	.906	.902	.866	.887	.874

	KJ2S	KJ3S	KJ4S	KJ5S	KJ6S	KJ7S	KJ8S	KJ9S	KJ10S
corr1	.742	.776	.760	.698	.761	.830	.777	.745	.744
corr2	.298	.326	.352	.265	.314	.290	.270	.280	.314
corr3	.672	.745	.728	.741	.739	.715	.728	.742	.744

	JF1S	JF2S	JF3S	JF4S	JF5S	JF6S	JF7S	JF8S	JF9S	JF10S
corr1	.715	.702	.675	.735	.712	.720	.708	.687	.734	.640
corr2	.298	.312	.270	.297	.297	.301	.275	.283	.287	.223
corr3	.765	.767	.794	.789	.769	.741	.732	.768	.821	.778

	JM1B	CH1B	CH2B	CH3B	CH4B	CH5B	CH6B	CH7B	CJ2B	CJ3B
corr1	.594	.711	.744	.735	.725	.682	.733	.716	.516	.482
corr2	.260	.280	.317	.298	.323	.292	.286	.294	.193	.214
corr3	.735	.681	.755	.717	.698	.689	.680	.687	.695	.686

	CJ4B	CJ5B	CJ6B	CJ7B	CJ8B	CJ9B	CJ10B	CJ11B	CJ12B	CJ13B
corr1	.434	.445	.435	.411	.461	.497	.473	.472	.471	.442
corr2	.151	.142	.150	.137	.170	.193	.171	.189	.195	.145
corr3	.646	.596	.648	.656	.684	.755	.762	.777	.802	.794

Table 13, continued

	JU01B	JU02B	JU03B	JU04B	JU05B	JU06B	JU07B	JU08B	JU09B	JU10B
corr1	.622	.595	.608	.627	.641	.622	.669	.673	.598	.616
corr2	.155	.132	.154	.192	.223	.157	.191	.134	.143	.165
corr3	.601	.572	.581	.650	.583	.578	.608	.536	.561	.563

3.2.2 Fit of random-walk model.

We turn now to the random-walk model already described by the qualitative axioms formulated in the theoretical section. The first and most common type of eye movement is stay, i.e., staying put in the same grid position. The second type of movement is moving forward according to the register-machine model. The third general motion is backtracking, i.e., the movement back to a position already occupied in the same column. We divide it into three kinds with relative frequency data for all three shown in Table 14. In particular, bktrck1, is the relative frequency of the backtrack from the second row to the top row. The motion labeled bktrck2 is a motion back to the top of the column but with two or more steps. Finally, bktrck3 is the motion from a row to the preceding row if the preceding row is not the top row. By breaking backtracking into these three mutually exclusive categories we cover almost all the cases and can have available the possibility of disentangling different kinds of motion. It might be asked, how does bktrck 2 occur in subtraction? It is to be remembered that the symbol that is placed within the grid where the point of regard is located is not always a digit; it can be an underline or an answer blank instead of one of the digits given in the exercise. The fourth type is skip, which is the movement which would be expected by the register machine model for the fixation after the correct fixation under study. Skipping in subtraction seems more frequent in Table 14 than it really is, because we counted motion from the second row to the answer space without stopping at the underline symbol as cases of skipping. Finally the fifth type of motion, other, includes all other movements not previously classified. The relative frequencies of these movements for various sessions and subjects are shown in Table 14.

Table 14

Relative Frequency Data for the Random-walk Model with Five Possible Movements

	KJ1A	KJ2A	KJ3A	KJ4A	KJ5A	KJ6A	KJ7A	KJ8A	KJ9A	KJ10A	KJ11A	KJ12A
stay	.673	.644	.688	.643	.687	.712	.704	.682	.583	.650	.665	.683
forward	.152	.174	.164	.110	.173	.161	.169	.182	.234	.191	.179	.173
bktrck1	.017	.025	.017	.011	.016	.012	.013	.012	.021	.019	.019	.016
bktrck2	.009	.005	.005	.005	.005	.003	.003	.003	.004	.004	.006	.003
bktrck3	.013	.012	.009	.008	.010	.007	.007	.020	.011	.014	.010	.009
skip	.028	.043	.038	.028	.037	.035	.031	.027	.051	.032	.039	.038
other	.105	.097	.076	.195	.071	.068	.072	.070	.093	.089	.081	.076

	JF2A	JF3A	JF4A	JF5A	JF6A	JF7A	JF8A	JF9A	JF10A	JF11A	JF12A
stay	.556	.589	.545	.584	.582	.602	.593	.598	.637	.569	.583
forward	.220	.220	.228	.219	.245	.183	.201	.216	.186	.219	.222
bktrck1	.028	.026	.025	.028	.022	.026	.033	.025	.028	.022	.030
bktrck2	.019	.009	.016	.004	.007	.003	.006	.006	.004	.010	.004
bktrck3	.023	.016	.022	.025	.016	.021	.025	.021	.024	.036	.031
skip	.058	.055	.070	.056	.059	.051	.060	.059	.044	.064	.050
other	.097	.085	.094	.084	.068	.113	.081	.075	.077	.080	.079

	KJ2S	KJ3S	KJ4S	KJ5S	KJ6S	KJ7S	KJ8S	KJ9S	KJ10S
stay	.589	.618	.564	.516	.580	.726	.656	.595	.629
forward	.120	.124	.150	.126	.141	.086	.094	.109	.117
bktrck1	.030	.030	.030	.065	.032	.024	.032	.044	.026
bktrck2	.007	.005	.009	.003	.005	.003	.004	.003	.008
bktrck3	.013	.015	.016	.016	.018	.011	.014	.014	.013
skip	.077	.073	.082	.087	.073	.037	.050	.065	.061
other	.164	.135	.151	.187	.152	.113	.150	.171	.145

Table 14, continued

	JF1S	JF2S	JF3S	JF4S	JF5S	JF6S	JF7S	JF8S	JF9S	JF10S
stay	.501	.461	.461	.522	.486	.502	.478	.454	.532	.475
forward	.180	.201	.188	.187	.202	.197	.190	.194	.168	.159
bktrck1	.039	.040	.045	.025	.028	.022	.034	.036	.025	.041
bktrck2	.011	.014	.024	.012	.015	.016	.013	.013	.014	.009
bktrck3	.018	.014	.010	.008	.014	.013	.019	.019	.018	.055
skip	.088	.103	.101	.101	.109	.101	.102	.105	.100	.082
other	.163	.167	.171	.145	.147	.150	.164	.179	.143	.180

	JM1B	CH1B	CH2B	CH3B	CH4B	CH5B	CH6B	CH7B	CJ2B	CJ3B
stay	.458	.599	.650	.631	.627	.573	.652	.624	.360	.328
forward	.112	.097	.092	.092	.100	.100	.080	.094	.097	.115
bktrck1	.052	.054	.056	.060	.056	.058	.036	.041	.070	.055
bktrck2	.012	.017	.009	.006	.007	.007	.007	.005	.015	.040
bktrck3	.071	.037	.028	.030	.025	.025	.025	.032	.058	.074
skip	.014	.216	.235	.258	.270	.023	.022	.032	.047	.043
other	.280	.174	.141	.154	.159	.214	.177	.171	.353	.344

	CJ4B	CJ5B	CJ6B	CJ7B	CJ8B	CJ9B	CJ10B	CJ11B	CJ12B	CJ13B
stay	.311	.319	.340	.309	.351	.376	.359	.337	.349	.339
forward	.095	.106	.113	.100	.101	.134	.100	.119	.112	.098
bktrck1	.054	.048	.076	.076	.081	.077	.081	.059	.076	.052
bktrck2	.026	.032	.014	.015	.018	.013	.020	.024	.020	.031
bktrck3	.076	.063	.061	.072	.060	.074	.089	.090	.071	.084
skip	.042	.038	.026	.031	.025	.016	.026	.022	.015	.025
other	.397	.394	.370	.397	.364	.310	.326	.349	.357	.370

Table 14, continued

	JU01B	JU02B	JU03B	JU04B	JU05B	JU06B	JU07B	JU08B	JU09B	JU10B
stay	.608	.612	.602	.595	.595	.586	.631	.657	.599	.601
forward	.046	.042	.047	.054	.064	.046	.055	.031	.045	.046
bktrck1	.028	.027	.036	.036	.038	.038	.032	.025	.036	.045
bktrck2	.013	.006	.010	.019	.008	.005	.003	.003	.004	.003
bktrcK3	.042	.048	.058	.059	.054	.045	.041	.034	.037	.042
skip	.012	.010	.008	.010	.010	.010	.013	.007	.010	.010
other	.251	.255	.240	.227	.230	.270	.226	.243	.270	.253

It is also of interest to determine whether or not the various steps in the random-walk model have different fixation durations associated with them. Inspection of Table 15 reveals no striking quantitative differences. On the other hand, for the two adult subjects (KJ and JP) the step bktrck2 has the minimum mean fixation duration in 32 of the 42 sessions and in 6 of 10 sessions for the younger student JU. For the other 20 sessions of the younger students (JM, CH, and CJ) the step skip is minimum in mean duration.

4 Discussion

The theory and data presented in this paper bear on a number of issues and problems that are appropriate to consider in conclusion.

To begin with, what new insight or information do eye movements give us about cognitive procedural models such as the register-machine model considered in detail in this article? There are four points we would make. First, the distributions of fixation durations, which suggest a nearly memoryless process, provide strong evidence that essential aspects of the information processing are stochastic in character. In this connection, it is worth recalling that the best current theories of randomness equate randomness with high complexity (the work of Kolmogorov, Martin-Löf, and others), and so to say that the processes are stochastic is to say that they are of high complexity. This being the case, they almost surely cannot even in principle be adequately represented in a deterministic fashion.

Second, the data show that even the well-trained subjects do not follow the register-machine model in detail. From Table 14 we can see that the only eye movements completely consistent with the register-machine

Table 15

Average Fixation Durations for Different Steps in the Random-walk Model

	JF2A	JF3A	JF4A	JF5A	JF6A	JF7A	JF8A	JF9A	JF10A	JF11A	JF12A
stay	.311	.292	.305	.270	.292	.236	.284	.283	.223	.343	.326
forward	.297	.253	.252	.246	.241	.239	.251	.255	.237	.276	.267
bktrck1	.278	.258	.270	.240	.255	.234	.258	.272	.212	.214	.291
bktrck2	.080	.102	.112	.127	.129	.204	.112	.132	.135	.112	.172
bktrck3	.305	.266	.251	.232	.275	.205	.256	.281	.222	.247	.292
skip	.304	.284	.245	.320	.292	.257	.276	.269	.249	.238	.292
other	.234	.231	.250	.200	.198	.169	.202	.218	.171	.219	.242

	KJ2S	KJ3S	KJ4S	KJ5S	KJ6S	KJ7S	KJ8S	KJ9S	KJ10S
stay	.263	.248	.288	.361	.283	.133	.187	.258	.255
forward	.271	.244	.267	.285	.237	.131	.185	.242	.229
bktrck1	.217	.231	.237	.404	.270	.111	.200	.220	.216
bktrck2	.154	.139	.138	.248	.140	.148	.109	.162	.126
bktrck3	.152	.148	.162	.210	.201	.084	.192	.176	.169
skip	.360	.281	.315	.375	.297	.196	.240	.303	.275
other	.174	.173	.189	.211	.183	.096	.136	.187	.174

	JF1S	JF2S	JF3S	JF4S	JF5S	JF6S	JF7S	JF8S	JF9S	JF10S
stay	.311	.325	.313	.338	.327	.318	.335	.338	.312	.261
forward	.246	.257	.232	.230	.226	.236	.247	.265	.218	.218
bktrck1	.277	.270	.331	.254	.274	.325	.247	.334	.296	.286
bktrck2	.190	.129	.145	.148	.123	.157	.139	.174	.123	.161
bktrck3	.259	.341	.354	.277	.307	.245	.269	.287	.319	.359
skip	.273	.286	.272	.303	.283	.277	.280	.280	.248	.244
other	.232	.238	.248	.270	.267	.268	.253	.260	.226	.210

Table 15, continued

	KJ1A	KJ2A	KJ3A	KJ4A	KJ5A	KJ6A	KJ7A	KJ8A	KJ9A	KJ10A	KJ11A	KJ12A
stay	.346	.344	.300	.232	.291	.260	.230	.234	.317	.333	.285	.281
forward	.437	.426	.363	.301	.333	.306	.286	.299	.404	.379	.345	.318
bktrck1	.159	.228	.176	.143	.198	.182	.139	.148	.188	.209	.170	.186
bktrck2	.120	.161	.138	.772	.111	.878	.150	.765	.115	.101	.712	.073
bktrck3	.238	.146	.182	.133	.195	.186	.135	.219	.148	.242	.173	.174
skip	.389	.394	.348	.261	.315	.265	.289	.261	.403	.334	.282	.306
other	.205	.158	.148	.181	.159	.128	.125	.142	.170	.201	.144	.129

	JM1B	CH1B	CH2B	CH3B	CH4B	CH5B	CH6B	CH7B	CJ2B	CJ3B
stay	.196	.156	.140	.152	.140	.152	.147	.151	.262	.231
forward	.224	.147	.140	.139	.157	.159	.145	.143	.246	.255
bktrck1	.249	.183	.186	.199	.194	.212	.208	.194	.326	.257
bktrck2	.197	.147	.143	.132	.127	.171	.134	.142	.280	.239
bktrck3	.188	.169	.154	.166	.168	.166	.125	.123	.322	.267
skip	.241	.106	.102	.092	.098	.100	.090	.083	.242	.247
other	.193	.139	.119	.117	.124	.133	.133	.128	.255	.230

	CJ4B	CJ5B	CJ6B	CJ7B	CJ8B	CJ9B	CJ10B	CJ11B	CJ12B	CJ13B
stay	.227	.281	.270	.236	.226	.271	.247	.242	.230	.232
forward	.239	.237	.248	.239	.236	.254	.258	.251	.259	.249
bktrck1	.278	.287	.321	.301	.291	.308	.263	.363	.276	.375
bktrck2	.190	.280	.218	.262	.242	.285	.218	.219	.222	.230
bktrck3	.266	.288	.358	.337	.303	.311	.283	.266	.282	.257
skip	.217	.233	.225	.192	.214	.172	.225	.273	.204	.268
other	.241	.268	.263	.255	.234	.273	.254	.232	.248	.271

Table 15, continued

	JU01B	JU02B	JU03B	JU04B	JU05B	JU06B	JU07B	JU08B	JU09B	JU10B
stay	.153	.131	.141	.159	.144	.129	.144	.132	.159	.161
forward	.183	.178	.169	.229	.205	.174	.196	.138	.177	.209
bktrck1	.142	.149	.119	.159	.131	.190	.153	.151	.184	.182
bktrck2	.153	.111	.110	.130	.119	.112	.071	.133	.163	.171
bktrck3	.174	.200	.213	.195	.189	.190	.164	.204	.282	.185
skip	.181	.177	.216	.225	.247	.150	.169	.143	.148	.120
other	.160	.146	.153	.173	.155	.152	.164	.132	.193	.168

model are those of stay and forward. Similar results are also to be found in the study of the correlation of the actual eye-movement path and the normative path of the register-machine model. A deeper and more difficult question is that of how the register-machine model should be revised to produce a more sophisticated normative model. This kind of question we are not able to pursue in any detail within the framework of data available to us in this article, but it is significant that we do not really know whether the eye movements classified as other play an important role in maintaining the efficiency of the subject as he moves from one digit to another and from one exercise to another. It is also not clear to what extent skipping should be encouraged and here again there is undoubtedly a tradeoff between reliability and speed, and the decision on this tradeoff would depend on the purposes for which the algorithms were being performed. The perfection of such a detailed normative model is probably not critical in the experimental area studied here, but the concept of perfecting performance algorithms that involve in an essential way eye movements is an important topic that seems to have received as yet little attention except in studies of reading. Even in this area, quantitative normative models have not been developed to any extent (see, e.g., the five excellent articles on reading in Monty & Senders, 1976).

Third, the eye-movement data show unequivocally that the perceptual component of the register-machine model is far too simple. Undoubtedly, if we were able to make the same kind of observations that we have made of eye movements of the internal processing, we would come to the same conclusion about the cognitive procedural aspects of the model as well.

Fourth, the revised and extended register-machine model formulated in terms of the qualitative axioms given in the first part of this article do fit the data at a relatively satisfactory level. We believe that the results are encouraging enough to warrant further studies in the same direction.

We now turn to some other considerations. It is apparent, both from theoretical and experimental viewpoints, that more work is needed to have an explicit identification of steps in the algorithm in relation to eye movements and processing time. To get such an identification, for example to estimate the process time and the associated eye movements with each individual step in the proposed algorithm, more specific model-theoretic assumptions must be made. Especially from the standpoint of processing time it is easy to make such further assumptions. Steps in this direction were already taken in Suppes (1973), but we are not yet satisfied with how this should be done to incorporate eye movements as well. It is also apparent that different highly specific models can be created, with different steps in detail, but it is also a problem to conceptualize these variants properly.

Another point is that for problems and models of the kind studied in this article it is too easy to think of the computer serving as an ideal version of a human subject. We want to emphasize how much our own view is distinct from that. The highly stochastic character of eye movements alone is at great variance from any current computer conceptions of perception. Our view is rather that it would be useful to try to build a computer model that more closely simulated what a human subject does than conversely.

From another standpoint, it is important to investigate conceptually different models for the processes we have studied here. An obvious and direct criticism of the register-machine model is the use of coordinates and the reference to grid locations. It is apparent that in detail this is not a realistic way of discussing human perception but a mathematical convenience that also must be regarded as a psychological fiction. What would be appropriate and more interesting would be to incorporate a geometry of symbols in the two-dimensional case and a geometry of objects in the three-dimensional case. The foundations of geometry in either of these instances is as yet far from satisfactory, and consequently fundamental work at a geometrical level is also required in order to create what we think would be sounder and more realistic models.

It is characteristic of the theoretical work we have pursued in this article that many of the details of eye movement have been ignored. For example, we have made no study of velocities or accelerations and it is important to know what better understanding of processing algorithms would be gained from a better understanding and study of these phenomena. Finally, in the same spirit we would remark that our current conception of the process of performing algorithms seems much too discrete. The basic register-machine model is a discrete model and yet the process, from a psychological standpoint, seems at a fundamental level to be more properly continuous than discrete. We have of course converted the register-machine model into a continuous-time stochastic process, as for example in the random-walk model we have already considered, but the processing steps remain discrete. Whether or not these simple steps also need to be replaced by a continuous version remains to be seen.

References

Anliker, J., Floyd, R., Morf, M., & Kailath, T. Biocybernetic factors in human perception and memory. Final report for the Advanced Research Projects Agency of the Department of Defense, ARPA Order Number 3177, July 1977. Stanford University, Center for Systems Research.

Cornsweet, T. N., & Crane, H. D. Accurate two-dimensional eyetracker using first and fourth Purkinje images. Journal of the Optical Society of America, 1973, 63, 921-918.

Crane, H. D., & Steele, C. M. An accurate three-dimensional eyetracker. Stanford Research Institute, unpublished report.

Ditchburn, R. W. Small involuntary eye movements: Solved and unsolved problems. In D. F. Fisher, R. A. Montz, & J. W. Senders (Eds.), Eye movements: Cognition and visual perception. Hillsdale, N.J.: Erlbaum, 1981. Pp. 227-235.

Groner, R. Hypothesen im Denkprozess. Bern, Stuttgart, Wien: Huber, 1978.

Just, M. A., & Carpenter, P. A. A theory of reading: From eye fixations to comprehension. Psychological Review, 1980, 87, 329-354.

Kraemer, H. C. The average error of a learning model, estimation and use in testing the fit of models. Psychometrika, 1965, 30, 343-352.

Monty, R. A., & Senders, J. W. (Eds.). Eye movements and psychological processes. Hillsdale, N.J.: Erlbaum, 1976.

Shepherdson, J. C., & Sturgis, H. E. The computability of partial recursive functions. Journal of the Association of Computing Machinery, 1963, 10, 217-255.

Suppes, P. Stimulus-response theory of finite automata. Journal of Mathematical Psychology, 1969, 6, 327-355.

Suppes, P. Facts and fantasies of Education. Phi Delta Kappa Monograph, 1973. Reprinted in M. C. Wittrock (Ed.), Changing education: Alternatives from educational research. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Suppes, P. Procedural semantics. In R. Haller & W. Grassl (Eds.), Language, logic, and philosophy. (Proceedings of the 4th International Wittgenstein Symposium, Kirchberg am Wechsel, Austria 1979.) Vienna: Hölder-Pichler-Tempsky, 1980. Pp. 27-35.

Suppes, P., Cohen, M., Laddaga, R., Anliker, J., & Floyd, R. Research on eye movements in arithmetic performance. In R. Groner & P. Fraisse (Eds.), Cognition and eye movement. Amsterdam: North-Holland (in press).

Suppes, P., Jerman, J., & Brian, D. Computer-assisted instruction: Stanford's 1965-66 arithmetic program. New York: Academic Press, 1968.

Suppes, P., & Morningstar, M. Computer-assisted instruction at Stanford, 1966-68: Data, models and evaluation of the arithmetic programs. Academic Press, 1972.

Appendix A

List of Subject Trials

The following is a list of child-subjects who passed initial screening for obvious features (e.g. seriously drooping eyelids) that would prevent the tracker from working. They were each brought to Stanford for at least two sessions with the eye-tracking system. We list the months in which we worked with them, sex, grade, and the disposition of the sessions. All the eighth-graders were from the Ravenswood School District, all the fifth-graders were from the Palo Alto Unified School District, and the third-graders were children of IMSSS personnel.

November 1979:

M.U.	boy, third grade	couldn't track

January 1980:

O.G.	girl, eighth grade	couldn't track
R.F.	girl, eighth grade	quit
G.H.	girl, eighth grade	dental infection

February through May 1980:

Y.D.	girl, eighth grade	couldn't track
J.M.	girl, eighth grade	difficult to track, 1 session completed
C.J.	girl, eighth grade	difficult to track

June 1980:

A.A.	boy, third grade	couldn't track

July through August 1980:

C.J.	girl, eighth grade	10 additional sessions completed
C.H.	boy, fifth grade	7 sessions completed
G.M.	boy, fifth grade	couldn't track
J.P.	boy, fifth grade	10 sessions completed

Appendix B

List of Arithmetic Structures

What follows is a description of the specifications defining the 20 types of children's exercises necessary for their generation. Note: a,b,c are single digits; if leftmost then they are strictly greater than zero. The number s_i is the sum of column i, not just the answer digit.

Type 1:

$$\begin{array}{rr} a & x \\ b & x \\ +\,c & x \\ \hline s_2 & s_1 \end{array} \qquad s_1 = 0,\ 2 < s_2 < 10$$

Type 2:

$$\begin{array}{rr} a & x \\ + & b \\ \hline s_2 & s_1 \end{array} \quad \text{or} \quad \begin{array}{rr} & b \\ +\,a & x \\ \hline s_2 & s_1 \end{array} \qquad 0 < s_1 < 10,\ 0 < s_2 < 10$$

Type 3:

$$\begin{array}{rr} a & x \\ +\,b & x \\ \hline s_2 & s_1 \end{array} \qquad 0 < s_1 < 10,\ 1 < s_2 < 10$$

Type 4:

$$\begin{array}{r} x \\ x \\ +\,x \\ \hline s_1 \end{array} \qquad 9 < s_1 < 20$$

Type 5:

$$\begin{array}{rr} a & x \\ +\,b & x \\ \hline s_2 & s_1 \end{array} \qquad 0 < s_1 < 10,\ 9 < s_2 < 19$$

Type 6:

$$\begin{array}{r} x \\ x \\ x \\ +\ x \\ \hline s_1 \end{array} \qquad 9 < s_1 < 20$$

Type 7:

$$\begin{array}{rr} a & x \\ +\ b & x \\ \hline s_2 & s_1 \end{array} \qquad 9 < s_1 < 19,\ 1 < s_2 < 9$$

Type 8:

$$\begin{array}{rr} a & x \\ +\ b & x \\ \hline s_2 & s_1 \end{array} \qquad 9 < s_1 < 19,\ 8 < s_2 < 19$$

Type 9:

$$\begin{array}{rrr} a & x & x \\ +\ b & x & x \\ \hline s_3 & s_2 & s_1 \end{array} \qquad 9 < s_1 < 19,\ 0 \leq s_2 < 9,\ 9 < s_3 < 19$$

Type 10:

$$\begin{array}{rr} a & x \\ b & x \\ + & c \\ \hline s_2 & s_1 \end{array} \quad \text{or} \quad \begin{array}{rr} a & x \\ & c \\ +\ b & x \\ \hline s_2 & s_1 \end{array} \quad \text{or} \quad \begin{array}{rr} & c \\ a & x \\ +\ b & x \\ \hline s_2 & s_1 \end{array} \qquad 9 < s_1 < 28,\ 8 < s_2 < 19$$

Type 11:

$$\begin{array}{rrr} a & x & x \\ +\ b & x & x \\ \hline s_3 & s_2 & s_1 \end{array} \qquad 0 < s_1 < 10,\ 9 < s_2 < 19,\ 8 < s_3 < 19$$

Type 12:

$$\begin{array}{rrr} a & x & x \\ +\ b & x & x \\ \hline s_3 & s_2 & s_1 \end{array} \qquad 9 < s_1 < 19,\ 8 < s_2 < 19,\ 1 < s_3 < 9$$

Type 13: $\begin{array}{rcc} & c & a \\ - & d & b \\ \hline \end{array}$ $\quad c > d > 0, \quad a \geq b > 0$

Type 14: $\begin{array}{rccc} & 1 & a & f \\ - & & b & g \\ \hline \end{array}$ $\quad a \geq b > 0, \quad f \geq g \geq 0$

Type 15: $\begin{array}{rccc} & 1 & a & f \\ - & 1 & b & g \\ \hline \end{array}$ $\quad a \geq b \geq 0, \quad f \geq g \geq 0$

Type 16: $\begin{array}{rcc} & c & a \\ - & d & b \\ \hline \end{array}$ $\quad c > d > 0, \quad 0 \leq a < b$

Type 17: $\begin{array}{rccc} & c & a & f \\ - & d & b & g \\ \hline \end{array}$ $\quad c \geq d > 0, \quad a \geq b > 0, \quad f \geq g \geq 0$

Type 18: $\begin{array}{rccc} & 1 & a & f \\ - & 1 & b & g \\ \hline \end{array}$ $\quad a > b \geq 0, \quad 0 \leq f < g$

Type 19: $\begin{array}{rccc} & c & a & f \\ - & d & b & g \\ \hline \end{array}$ $\quad c > d > 0, \quad 0 \leq a < b, \quad f \geq g \geq 0$

Type 20: $\begin{array}{rccc} & c & a & f \\ - & d & b & g \\ \hline \end{array}$ $\quad c > d > 0, \quad 0 \leq a \leq b, \quad 0 \leq f < g$

Appendix C

Register-machine Models for Addition and Subtraction

We give in this appendix the full register-machine models for general column addition, i.e., for exercises having an arbitrary number of rows and columns, and also the full subtraction model. For this purpose, it is convenient to add to the instructions given at the beginning of the article the following four:

JUMP L: Jump to line labeled L.

END: Terminate processing of current exercise.

PLUS [Reg]: Add 1 to the contents of register [Reg.].

MINUS [Reg.]: Subtract 1 from the contents of register [Reg].

The meaning of each of these four additional instructions is apparent. It is obvious that, for example, we do not from a formal standpoint need a JUMP instruction. We can write the register-machine programs with the conditional JUMP instruction but it is convenient and simple to have the unconditional instruction as well. We also need to extend the nonelementary LOOKUP to include subtraction as well as addition. The instruction is interpreted to subtract the contents of register [Reg2] from the contents of [Reg1] and then store the result in [Reg1]. These facts include storing a minus sign with the result if the result of the subtraction is negative.

For the general case of column addition it is also convenient to have two subroutines, one for vertical scanning of the left-hand side for irregular rows--that is, the exercise is not simply a rectangular array with each row having exactly the same number of columns--the other for outputting. These two subroutines are the following:

V-SCAN SUBROUTINE

```
1.  rdv     READIN*
2.          JUMP(0-9,-) SS,finv*
3.          ATTEND(+1,-1)*
4.          READIN*
5.          JUMP(-) SS,finv*
6.          ATTEND(+0,+1)*
7.          JUMP rdv*
8.  finv    EXIT*
```

The OUTPUT subroutine is the following:

OUTPUT SUBROUTINE

```
1.  put     OUTRIGHT NSS*
2.          DELETERIGHT NSS*
3.          ATTEND(0,+1)*SS*
4.          JUMP( ) NSS,fino*
5.          JUMP put*
6.  fino    exit*
```

The full model for column addition is then the following:

COLUMN ADDITION MODEL

```
1.          ATTEND(1,1)
2.          READIN
3.          COPY SS in OP*
4.          ATTEND(+1,+0)
5.          READIN
6.  OPR     LOOKUP OP + SS
7.  rd      ATTEND (+1,0)
```

```
8.              READIN
9.              JUMP(0-9) SS,opr
10.             JUMP( ) SS,rd
11.             ATTEND(+1,0)
12.             OUTRIGHT OP
13.             DELETERIGHT OP
14.             COPY OP in NSS
15.             ATTEND(1,+1)
16.             V-SCAN
17.             JUMP(-) SS,fin
18.  car        COPY NSS in OP*
19.             JUMP opr
20.  fin        JUMP ( ) NSS,out
21.             ATTEND(+1,+1)
22.             OUTPUT(NSS)
23.  out        END
```

Finally, the full subtraction model is the following:

SUBTRACTION MODEL

```
1           ATTEND(1,1)
2           READIN :38
3           COPY 0 in NSS
4           COPY SS in OP
5           ATTEND(+1,+0)
6           READIN
7   opr     LOOKUP OP - SS
8           ATTEND(+1,+0)
9           OUTRIGHT OP
10          DELETERIGHT OP
11          ATTEND(1,+1)
12          READIN
13          JUMP( ) SS,fin
14          JUMP(-) OP,bor
15  on      COPY SS in OP
16          ATTEND(+1,+0)
17          READIN
18          JUMP( )SS,easy
19          JUMP opr
20  easy    ATTEND(+1,+0)
21          OUTRIGHT OP
22          DELETERIGHT OP
23  again   ATTEND(1,+1)
24          READIN
25          JUMP( ) SS,fin
26          ATTEND(+2,+0)
27          OUTRIGHT SS
28          DELETERIGHT SS
29          JUMP again
30  bor     JUMP(0) SS,over
31          MINUS SS
32          JUMP on
33  over    PLUS NSS
34          ATTEND(1,+1)
35          READIN
36          JUMP(0) SS, over
37          ATTEND(1, -NSS)
38          READIN
39  line    COPY SS in OP
40          ATTEND(+1,+0)
41          READIN
42          COPY 9 in OP
43          JUMP( ) SS,easya
44          LOOKUP OP + SS
45          ATTEND(+1,+0)
46          OUTRIGHT OP
```

```
        DELETERIGHT OP
        MINUS NSS
        ATTEND (1,+1)
        READIN
        JUMP(0) SS,line
        MINUS SS
        JUMP on
fin     END
easya   ATTEND(+1,+9)
        OUTRIGHT OP
        DELETERIGHT OP
againa  ATTEND(1,+1)
        READIN
        JUMP( ) SS,fin
        JUMP(0) NSS,outa
        JUMP(0) SS,linea
        MINUS SS
        COPY 0 in NSS
outa    ATTEND(+2,+0)
        OUTRIGHT SS
        DELETERIGHT SS
        JUMP againa
linea   COPY 9 in OP
        ATTEND(+2,+0)
        OUTRIGHT OP
        DELETERIGHT OP
        JUMP againa
```

We will now step through the subtraction model with a specific exercise so that the reader can obtain more insight into how one of the models actually works. The exercise that will be performed is 1073-82. The grid positions are

1073	1,4	1,3	1,2	1,1
82			2,2	2,1
	3,4	3,3	3,2	3,1

Hopefully when we are done the contents of (3,1) will be 1, of (3,2) will be 9, of (3,3) will be 9, and of (3,4) will be 0.

We first ATTEND (1,1). Then we read 3 into SS. We COPY 0 in NSS. We COPY 3 into OP. Then we ATTEND (2,1) which has a 2. We read 2 into SS.

Next we LOOKUP 3 - 2. This places a 1 in OP. We now ATTEND (3,1). This has a blank. We now write 1 in the grid position (3,1). Then we delete the 1 in OP. Now we have finished the first column and are ready to move to the second. We ATTEND (1,2) which has a 7 and read the 7 into SS. Since SS is not blank we do not JUMP to line FIN. Since OP does not have a - in it we do not JUMP to line BOR. We COPY 7 in to OP. We ATTEND (2,2) and READIN the 8 into SS. Since SS is not blank we do not go to line EASY. We JUMP to line OPR. We LOOKUP 7 - 8. This puts a -9 in OP. We now ATTEND (3,2) which is blank. We write the rightmost character of OP, which is 9, in (3,2) and then delete the 9, leaving - in OP.

The information about the borrow is preserved as we move to the next column. We ATTEND (1,3) which has a 0 in it, and read 0 into SS. Since SS is not blank we do not JUMP to FIN, and since OP does have - in it we jump to line BOR. We now enter the part of the subtraction algorithm devoted to borrowing. Since SS does have a 0 in it we JUMP to line OVER which signifies that we cannot borrow yet since the column immediately to the left has a zero in it. We now add 1 to NSS, which will keep track of how many column shifts we have to make before finding a column with a nonzero entry. We ATTEND (1,4) and READIN 1 into SS. Since SS is no longer zero we do not need to search any more and we are ready to continue programming. We ATTEND (1,3) and read 0 into SS. We COPY 0 into OP. We then ATTEND (2,3) and read the blank into SS. We COPY 9 in OP. Then we have a blank in SS so we JUMP to line EASYA. Lines EASY and EASYA begin two sections of the algorithm that correspond to the processes needed to complete a subtraction exercise when there are no longer any more digits in the subtrahend. The section starting with line EASY corresponds to the case when we encounter this situation when we are not borrowing, and EASYA starts the section when we are. We ATTEND (2,4) and write 9 at this position. We delete the 9 in OP and ATTEND (1,4). We read the 1 at (1,4) into SS and since SS is not blank we do not JUMP to line FIN. Since neither SS nor NSS has a zero in it we do no jumping and we reduce SS by one to zero. We COPY zero in NSS. We ATTEND (3,4) and write 0 in that grid position. We delete the zero in SS so that SS now contains a blank. We JUMP to line AGAINA. We ATTEND(1,5). We READIN the blank into SS. Since SS is now blank we JUMP to FIN which ENDs the exercise.

Appendix D

Professional Project Personnel

1. Dr. Patrick Suppes, Director, IMSSS. Principal Investigator.

2. Robert Laddaga, Research Assistant, IMSSS. Project Director.

3. Michael Cohen, Graduate Research Assistant, IMSSS and Department of Statistics.

4. Dr. James Anliker, Senior Research Associate, IMSSS.

5. Robert Floyd, Senior Systems Programmer.

Appendix E

List of Lectures

Lectures by Mr. Laddaga:

Research on Process Models of Basic Arithmetic Skills: Status Report. National Science Foundation, NSF RISE Project Director's Conference, Stanford University, Stanford, California, xxx 1978.

Process Models of Basic Arithmetic Skills. National Institute of Education-National Science Foundation Joint Program for Research on Cognitive Process and the Structure of Knowledge in Science and Mathematics NSF, November, 1979.

Lectures by Dr. Suppes:

September 1, 1979	Procedural Semantics. Fourth International Wittgenstein Symposium, Kirchberg, Austria
September 4, 1979	Data and Theory on Eye Movements in Performance of Algorithms. E. L. Thorndike Award Lecture, American Psychological Association, New York City
October 11, 1979	Some Reseach Issues in Computer-assisted Instruction. 1979 Award Lecture in honor of S. Richard Silverman. Central Institute for the Deaf, St. Louis, Missouri
November 15, 1979	Procedural Semantics: Philosophical and Psychological Aspects. Aix-en-Provence, France
August 14, 1980	A Procedural Approach Toward Mathematics Education. ICME IV, University of California, Berkeley